Rockefeller Series in Science and Technology 3

Alessandro N. Vargas, Leonardo Acho, Gisela Pujol *Editors*

Huang Xiao

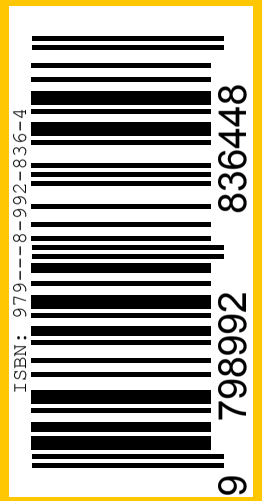
Generative AI as the "Silly Language Partner"

Human–Machine Interaction in Junior High CSL Education

ISSN 3067-0667

ROCKEFELLER

NEW YORK, NY


ISBN: 979-8-992-836-4
9 798992 836448

Rockefeller Series in Science and Technology

Open Access • Book Series

Volume 3

Series Editor

Alessandro N. Vargas • UTFPR Brazil • UC Berkeley USA

For further volumes:
https://www.rockefellerpub.com/series

The Rockefeller Series in Science and Technology (ISSN 3067-0667) publishes research monographs, edited works, collections of papers, reviews of previously published studies, and advanced textbooks covering all relevant subjects related to science and technology.

Monographs of this series can contain the following contributions:

- Book chapters that revisit and expand the interpretation of previously published works;
- Timely discussions of a relevant topic;
- Literature review with insights and novel interpretations;
- A carefully designed study that sparks interest in the research community;
- Research tailored for graduate students and professionals;
- Technical reports;
- Other topics related to the book series.

The monographs published in the Rockefeller Series on Science and Technology attract the interest of researchers, students, and professionals.

Rockefeller Publishing Co. welcomes book ideas. Potential authors who wish to submit a book proposal should contact the Editors at `https://rockefellerpub.com/books`.

Generative AI as the "Silly Language Partner": Human–Machine Interaction in Junior High CSL Education

Huang Xiao • Krirk University, Bangkok, Thailand

Editor(s)

Alessandro N. Vargas • UTFPR Brazil • UC Berkeley USA
Leonardo Acho • UPC BarcelonaTech Spain
Gisela Pujol • UPC BarcelonaTech Spain

Huang Xiao
Krirk University, Thailand
3 Soi Ramindra 1, Khwaeng Anusawari,
Khet Bang Khen, Bangkok 10220
huang.xiao@krirk.ac.th

ISSN 3067-0667
Rockefeller Series in Science and Technology

ISBN 979-8-9928364-4-8 (eBook)
DOI: https://doi.org/10.63408/979-8-9928364-4-8

© The Author(s) 2026. This book is an open access publication.

Notice. This book is published under the Creative Commons Attribution 4.0 International License (CC BY 4.0). This license allows authors to publish their work openly, enabling others to use, share, adapt, distribute, and reproduce the content in any medium or format, as long as appropriate credit is given to the original author(s) and source.

The publisher, authors, and editors believe that the information presented in this book is accurate as of the publication date. However, they do not offer any guarantee, express or implied, regarding the accuracy or completeness of the material contained herein, nor do they assume responsibility for any errors or omissions. The application of the information in this book should be taken at the reader's discretion.

The publisher remains neutral about jurisdictional claims in published maps and institutional affiliations.

This open-access book is published and distributed online by Rockefeller Publishing Co.

Rockefeller Publishing Co., 45 Rockefeller Plaza 20th Floor, New York, NY 10111 USA.

Contents

Generative AI as the "Silly Language Partner": Human–Machine Interaction in Junior High CSL Education i
Huang Xiao • Krirk University, Bangkok, Thailand

Abstract xiii

Author Biography xv

1 Introduction 1
1 Research Background 1
1.1 A New Era of AI Technology in Language Education . . . 1
1.2 The Specific Challenges Facing Junior High Second Language Learners 2
1.3 The Classroom Context and Curriculum Design of This Study 3
2 Research Purpose and Research Questions 4
2.1 Research Purpose 4
2.2 Research Questions 4
3 Significance of the Study 5
3.1 Theoretical Significance 5
3.2 Practical Significance 5
3.3 Methodological Significance 6
4 Definition of Core Concepts 7
5 Book Structure 8
6 Reader Navigation: How This Book Answers Its Research Questions 9
7 Conceptual Framework: How Research Questions, Theories, and Findings Connect 9
7.1 Theoretical Frameworks and Their Corresponding Research Questions 9
7.2 From Theories to Findings to Theoretical Contributions . 10

7.3 Summary Table: Research Questions, Theories, and Chapters 11
8 Chapter Summary 11

2 Literature Review 13
1 Introduction 13
2 Second Language Acquisition Theories: Interaction, Output, and Affect 14
2.1 The Interaction Hypothesis: Negotiation of Meaning and Language Acquisition 14
2.2 The Output Hypothesis: Noticing Trigger, Hypothesis Testing, and Metalinguistic Reflection 15
2.3 The Affective Filter Hypothesis: Anxiety, Motivation, and Language Acquisition 16
2.4 Implications of the Theoretical Review for This Study . . . 16
3 Computer-Assisted Language Learning and AI Roles 17
3.1 The Trajectory of CALL: From Behaviorism to Generative AI 17
3.2 Four Roles of AI in Education 18
3.3 Extensions and Debates on the AI Role Framework 19
3.4 Implications of the Theoretical Review for This Study . . . 19
4 Privacy and Ethics in Educational AI 20
4.1 The Responsible Research and Innovation Framework . . . 20
4.2 Privacy Principles for Educational AI 21
4.3 Parental Concerns About Educational AI 22
4.4 Implications of the Theoretical Review for This Study . . . 23
5 Gaps in Existing Research and the Positioning of This Study 23
5.1 Major Gaps in Existing Research 23
5.2 Theoretical Positioning of This Study 24
5.3 Relationship Between This Study and Existing Theories . 25
6 Chapter Summary 25

3 Research Methodology 27
1 Introduction 27
2 Research Paradigm and Design 28
2.1 Rationale for Choosing the Qualitative Research Paradigm 28
2.2 Single-Case Longitudinal Design 29
2.3 Statement of the Researcher's Role 29
3 Research Field and Participants 30
3.1 School and Classroom Context 30
3.2 Student Participants 31
3.3 Parent Participants 31
4 Data Collection Methods 32
4.1 Classroom Observations and Video Recordings 32

4.2 Student-AI Dialogue Logs 32
4.3 Student Task Cards 33
4.4 Semi-Structured Interviews 33
4.5 Open-Ended Parent Communication 34
4.6 Teacher Reflective Journals 35
5 Data Collection Timeline 35
6 Data Analysis Methods 36
6.1 Thematic Analysis 36
6.2 Analytical Framework: Theory-Data Dialogue 38
6.3 Analysis Software 38
7 Trustworthiness of the Study 39
7.1 Credibility 39
7.2 Transferability 40
7.3 Dependability 41
7.4 Confirmability 41
8 Researcher's Reflexivity 41
8.1 Managing the Teacher-Researcher Dual Role 42
8.2 Researcher's Pre-understandings and Assumptions 42
8.3 Managing Emotional Responses 42
9 Ethical Considerations 43
9.1 Informed Consent 43
9.2 Privacy and Data Protection 43
9.3 Minimizing Harm and Benefits 44
10 Chapter Summary 45

4 Research Findings I – Human-Machine Negotiation of Meaning 47
1 Introduction 47
2 Theme 1: "It Doesn't Understand" - Triggers of Meaning Negotiation and Responses 48
2.1 The Al's "Failure to Understand" as a Trigger for Negotiation 48
2.2 Initial Stage (Weeks 1-4): Tension, Silence, and Tentativeness 49
2.3 Middle Stage (Weeks 5-9): From Tension to Active Correction 49
2.4 Later Stage (Weeks 13-18): Calm, Strategic Responses . . 50
2.5 The Transition Trajectory from "Surprise" to "Expectation" 51
3 Theme 2: "Let Me Teach You" - Students' Meaning Negotiation Strategies as "Correctors" 52
3.1 Types of Meaning Negotiation Strategies 52
3.2 Relationship Between Strategy Choice and Language Proficiency 53
3.3 "Teaching the AI" as a Unique Negotiation Context 54
3.4 Effectiveness and Failure of Strategies 55

4 Theme 3: "The Process of Teaching It Is Also Learning" - The Impact on Language Output 56
4.1 Noticing Trigger: How Students "Notice" Their Own Language Problems 56
4.2 Hypothesis Testing: How Students Use Output to "Test" Language Hypotheses 57
4.3 Output Modification: How Students Adjust Their Own Language 58
4.4 Students' Metacognitive Development: "The Process of Teaching It Is Also Learning" 59
5 Chapter Summary 61

5 Research Findings II – The Fluidity and Evolution of AI Roles 63
1 Introduction 63
1.1 Theme 1: "What Is It Like?" - Students' Diverse Perceptions of AI Roles 64
1.2 Emergence of Role Types 64
1.3 Relationship Between Role Perception and Language Proficiency 66
1.4 Relationship Between Role Perception and Personality Traits 67
2 Theme 2: "It Changes" - The Fluidity and Context-Dependence of AI Roles 68
2.1 Role Switching Within the Same Student, Same Lesson . 68
2.2 Key Events Triggering Role Switching 69
2.3 Pedagogical Implications of "Role Fluidity" 69
3 Theme 3: "From Opponent to Tool" - The Semester-Long Evolution of Al Role Perception 70
3.1 S7's Evolution Trajectory: From "Testing Machine" to "Learning Partner" 70
3.2 S3's Evolution Trajectory: From "Peer in Need" to "Stable Peer" 71
3.3 S11's Evolution Trajectory: From "Tool" to "More Efficient Tool" 72
3.4 Comparison of Three Evolution Trajectories 73
4 Chapter Summary 74

6 Research Findings III - Privacy, Ethics, and Parental Communication 77
1 Introduction 77
2 Theme 1: "What Can I Do?" - The Teacher's Ethical Practices and Decisions 78
2.1 Ethical Preparation at the Beginning of the Semester: Parent Letter and Informed Consent 78
2.2 Privacy Protection Measures and Implementation in the Classroom 79

2.3 Handling Unexpected Ethical Dilemmas 80
2.4 Ethical Reflections on the Teacher-Researcher Dual Role 81
2.5 The Teacher's "Do Not Do" List 82
3 Theme 2: "What Am I Worried About?" - Parents' Narratives of Concern 82
3.1 Main Concerns at the Beginning of the Semester 82
3.2 Evolution of Concerns: From Beginning to End of Semester 84
3.3 Unresolved Concerns 86
4 Theme 3: "I See Changes" - Parents' Acceptance and Recognition 87
4.1 Changes Parents Observed in Their Children 87
4.2 Parents' Reassessment of Al's Value 88
4.3 Parents' Suggestions for Future AI Use 88
5 Chapter Summary 89

7 Discussion and Conclusion 91
1 Summary of Research Findings 91
1.1 Answer to RQ1: Human-Machine Negotiation of Meaning 91
1.2 Answer to RQ2: AI Role Perception 92
1.3 Answer to RQ3: Privacy, Ethics, and Digital Citizenship Education 92
2 Dialogue with Theory 93
2.1 Extending the Interaction Hypothesis and Output Hypothesis 93
2.2 Re-examining the Affective Filter Hypothesis 94
2.3 Extending the AI Role Framework 94
2.4 Applying the RRI Framework to Educational AI 95
3 Theoretical Contributions 96
3.1 Conceptualizing "Human-Machine Negotiation of Meaning" 96
3.2 Proposing the Concept of "Role Fluidity" 97
3.3 The "From Opponent to Partner" Evolution Trajectory Model 97
3.4 An Operationalization Framework for "Teacher Ethical Practices" 98
4 Pedagogical Recommendations 98
4.1 Regarding AI "Personality" Design 98
4.2 Regarding Activity Design 99
4.3 Regarding Privacy and Ethics 99
4.4 Regarding Emotional Support 100
4.5 BOX: Quick Reference - 5 Things Teachers Can Do Tomorrow 100
5 Research Limitations 101
5.1 Methodological Limitations 101

5.2 Technical Limitations 101
5.3 Theoretical Limitations 102
6 Future Research Directions 102
6.1 Comparative Research 102
6.2 Longitudinal Research 103
6.3 Design Research 103
6.4 Ethics and Policy Research 103
7 Conclusion 104

Foreword

The origin of this book can be traced back to a small moment I observed in my classroom in the autumn of 2024. That year, I was conducting a long-term research project in an international school.

It was an ordinary afternoon when my 14 junior high students were engaged in dialogue practice with ChatGPT. According to the curriculum design for that week, the AI had been configured as a "directionally challenged little fool"—it consistently mistook "turn left" for "turn right" and "go straight" for "turn." When S7 (a Korean boy who rarely dared to speak in class) loudly declared to his tablet, "No! Not left turn. Go straight first!"—the expression on his face, a mixture of surprise, excitement, and a sense of accomplishment, stayed with me for a long time.

Just a few weeks earlier, this same boy had remained silent in Chinese class, afraid of being laughed at by his peers for making mistakes. Yet here he was, "correcting" an "imperfect" AI—and thoroughly enjoying it.

That moment made me realize something: perhaps the greatest value of AI in the language classroom lies not in its "perfection," but precisely in its "imperfection." An AI that never makes mistakes is a perfect tool, but not necessarily a good "language partner." In contrast, an AI that makes mistakes, that sometimes "doesn't understand," that requires students to "teach" it—such an AI might create the kind of "role reversal" that is difficult to achieve in traditional classrooms. Students shift from "the corrected" to "the corrector," from "the taught" to "the teacher."

It was this intuition that gave birth to the research documented in this book.

In the fall semester of 2025, I conducted an 18-week qualitative study in my own junior high Chinese as a Second Language classroom at an international school. Fourteen students, seven different native language backgrounds, weekly AI-assisted lessons, and data collected throughout the semester—classroom observations, student interviews, parent questionnaires, and teacher reflective journals—form the analytical foundation of this book.

This book seeks to answer three interrelated questions:

First, the interactional dimension: What kinds of meaning negotiation behaviors occur between students and the AI? When the AI "fails to understand" or "makes a

mistake," how do students respond? How does the AI's "intentional error-making" design affect students' language output and modification?

Second, the perceptual dimension: How do students perceive the AI's role in their Chinese learning? How does this role perception evolve over time and across interactional situations?

Third, the ethical dimension: What privacy and ethical protection strategies does the teacher employ in AI-assisted teaching? What narratives of concern and acceptance do parents exhibit regarding the introduction of AI into the Chinese classroom?

This book is neither a simple "effectiveness evaluation" of AI nor an operational "how-to" manual. It seeks to present the complex, rich, and sometimes contradictory stories that unfold when AI enters a real classroom—the evolutionary trajectory of students from tense silence to excited correction to calm strategic use, the multiple role shifts (tool, opponent, game partner, peer in need, etc.) that the same student experiences within a single lesson, the transformation of parental attitudes from concern to conditional support, and the decisions, confusions, and reflections I experienced throughout this process as a teacher-researcher.

Methodologically, this book adopts a purely qualitative research paradigm. I chose "thick description" over "statistical significance," "contextualized understanding" over "decontextualized generalization." I am aware that one classroom, 14 students, and 18 weeks of data cannot yield statistically "representative" conclusions. But I believe that through sufficiently detailed description, the stories and insights in this book can resonate with other educators in similar contexts. When you read about S7's transformation from "not daring to speak" to "loudly correcting the AI," you may be reminded of a similar student in your own classroom. When you read about Parent P7's shift from "worrying about my child's voice being recorded" to "I trust that you are paying attention," you may reflect on how to communicate more effectively with parents in your own teaching context.

Another distinctive feature of this book is the researcher's dual role. I was the Chinese teacher for this class, the curriculum designer, and the researcher. This "teacher-researcher" dual identity is both an advantage (deep access to the field, trust-building, long-term observation) and a challenge (students might give answers "the teacher wants to hear," the researcher might have emotional investment in their own curriculum design). In Chapter 3, I discuss in detail the strategies for managing this dual role; throughout the research process, I worked to maintain trustworthiness through reflective journals, peer review, and active searching for disconfirming evidence.

This book is written for three primary audiences. First, CSL (Chinese as a Second Language) practitioners working in international school settings—particularly those at the junior high level—will find detailed, ready-to-adapt teaching materials (task cards, parent letters, observation forms) and pedagogical recommendations grounded in 18 weeks of real classroom experience. Second, second language acquisition (SLA) researchers interested in human-machine interaction will find qualitative data on meaning negotiation, output modification, and role perception that extend classic theories (Interaction Hypothesis, Output Hypothesis) into the AI era.

Third, educational policymakers and school administrators grappling with the responsible integration of generative AI will find a documented ethical practice framework, including parent communication strategies and privacy protection measures that can inform institutional policies.

Before outlining the book's structure, I want to name three theoretical concepts that emerge from this study and are developed fully in Chapter 7. These concepts constitute the book's most distinctive contribution to the literature: (a) *human-machine negotiation of meaning*, an extension of Long's Interaction Hypothesis to AI-mediated interaction; (b) *role fluidity*, a dynamic supplement to Luckin et al.'s (2016) AI role framework; and (c) the *opponent-to-partner evolution trajectory*, a model of how lower-proficiency learners' perceptions of AI can transform over time. Even as specific AI tools become outdated, these concepts—and the empirical phenomena they describe—are likely to retain their relevance.

The structure of this book is as follows: Chapter 1 (*Introduction*) presents the research background, questions, and significance. Chapter 2 (*Literature Review*) systematically reviews second language acquisition theories, computer-assisted language learning and AI role frameworks, and privacy and ethics research in educational AI. Chapter 3 (*Research Methodology*) details the research design, data collection, and analysis methods. Chapters 4 through 6 present the three core findings (human-machine negotiation of meaning, AI role perception, and privacy and ethics). Chapter 7 (*Discussion and Conclusion*) engages in theoretical dialogue, proposes pedagogical recommendations, and identifies limitations and future directions.

Additionally, the appendices contain a wealth of ready-to-use teaching materials—parent letter templates, student task card templates, interview protocols, classroom observation forms, teacher reflection log templates, coding schemes, and more. All of these materials are drawn from the 18 weeks of authentic teaching practice and can be adapted by other teachers to their own classroom contexts.

I am deeply grateful to the 14 students and their parents. The students participated in this research with sincerity and openness, sharing their genuine thoughts in interviews and allowing me to document their learning processes in class. Parents expressed their concerns and suggestions through questionnaires and interviews, helping me continuously improve my instructional design and communication strategies. Without their trust and participation, this book would not exist.

I am also grateful to the school leaders and colleagues. They supported this research, provided technical and administrative assistance for the use of AI tools, and offered encouragement and constructive feedback throughout the process.

Finally, I acknowledge the limitations of this book. This is a case study situated in a specific context—an international school, junior high level, HSK Level 3 proficiency, ChatGPT as the primary tool. These boundary conditions mean that the findings of this book should not and cannot be "directly applied" to other contexts. What I hope for is that other educators and researchers will find inspiration in these pages, testing, revising, or challenging these findings in their own contexts, collectively advancing the important dialogue on the integration of AI and language education.

Since 2023, generative AI technologies have entered the field of education at an unprecedented pace. Every day brings new tools, new application scenarios, and new research. In the face of such rapid technological iteration, the specific tools described in this book (a particular version of ChatGPT) and the specific activity designs may soon become outdated. But I hope that the core insights this book reveals—the value of meaning negotiation, the existence of role fluidity, the importance of ethical practices, and the unique mechanism of "teaching the AI" as learning—will have more lasting significance.

AI will not replace teachers. But teachers who know how to collaborate with AI may be better equipped to serve their students' growth. This book represents one exploration in that direction. If it inspires more dialogue, practice, and research, then my original intention will have been fulfilled.

Abstract

This study is a qualitative investigation of AI-assisted teaching in junior high Chinese as a Second Language (CSL) classrooms. Focusing on a junior high CSL classroom at an international school (14 students, approximately HSK Level 3 proficiency), this longitudinal study tracked an entire semester (18 weeks) to explore three interrelated core questions: human-machine negotiation of meaning, students' perceptions of AI roles and their evolution, and privacy and ethical practices in AI-assisted teaching.

The findings reveal: First, the AI's "intentional error-making" design successfully created abundant opportunities for meaning negotiation. Students' initial responses to the AI's "failure to understand" followed an evolutionary trajectory from "surprise" to "expectation" (tension-silence stage → excitement-correction stage → calm-strategic stage). Students employed five meaning negotiation strategies (repetition, paraphrase, simplification, code-switching to English, and multimodal assistance) and developed metacognitive understanding through the process of "teaching the AI" ("the process of teaching it is also learning").

Second, students' perceptions of AI roles were diverse (tool, peer in need, fallible opponent, coach, teacher, etc.) and fluid—the same student, in the same lesson, could assign different roles to the AI depending on changes in the interactional context. Students with lower language proficiency showed more dramatic evolution in role perception (S7's trajectory: testing machine → fallible opponent → game partner → learning partner), while high-proficiency students consistently perceived the AI as a tool.

Third, teachers need to adopt multi-layered, sustained ethical practices (parental informed consent at the semester's beginning, in-classroom privacy protection measures, timely handling of unexpected ethical dilemmas, and continuous reflexive reflection). Parents' concerns were multi-dimensional (data privacy, screen time, technology dependence, content appropriateness, and socio-emotional development), but through transparent communication and observing positive changes in their children, some parents' worries were transformed into support.

Theoretically, the study advances three interconnected concepts as extensions to existing frameworks. First, "human-machine negotiation of meaning" extends

Long's (1996) Interaction Hypothesis by highlighting the role of predictable AI errors in triggering output modification. Second, "role fluidity" challenges Luckin et al.'s (2016) assumption of fixed AI roles, demonstrating that the same student can assign different roles to the AI within a single lesson. Third, the "opponent-to-partner evolution trajectory" models how lower-proficiency learners' AI role perceptions transform over time (testing machine → fallible opponent → game partner → learning partner). These concepts are proposed for further testing and refinement in future research. Practically, this study provides actionable AI teaching strategies, an ethical practice framework, and parent communication guidelines for international school CSL teachers.

Keywords: AI-assisted language teaching; human-machine negotiation of meaning; AI role perception; role fluidity; intentional error design; teacher-researcher dual role; junior high Chinese as a Second Language; international school; multilingual learners; educational ethics; qualitative research.

Author Biography

Dr. Huang Xiao (b. 1988) is a scholar of contemporary art, visual culture, and interdisciplinary humanities. She received her PhD in Comparative Literature from Fudan University in China and subsequently served as a postdoctoral researcher and visiting scholar at Brigham Young University in the United States. Over the years, she has taught at various international institutions, including Adam Mickiewicz University (Poland), Universidad del Mar (Mexico), and the Technology University of Puebla (Mexico), etc.

Currently, Dr. Huang is a Professor at the International College of Art, Krirk University (Thailand), where she also serves as a doctoral and master's supervisor. At the same time, she is a consultant on Chinese language and culture at the newly established Turkey–Thailand Center in Bangkok, an academic expert reviewer for

the Media and Oral Communication Arts Research Centre at Huanggang Normal University (China), and an academic moderator of the Chinese Art Phenomenology Salon.

Dr. Huang is also a Contributing Author and Core Academic Committee Member of the *Journal of Art Spectrum*. She serves as a Guest Peer Reviewer for the *Asian Arts and Society Journal* and for the *Global Review of Humanities, Arts and Society*.

Her publications include the academic monographs *Between Ice and Melt: Embodied Perception, Cross-Cultural Emotion, and Artistic Transformation in Educational Practice*; *Three Chinese Contemporary Artists: Plural Dimensions, Plural Revelations*; the cross-disciplinary experimental art anthology *The Presence of the Transgressive Body: Art as Political Manifesto and Public Intervention*; the Chinese translation of *A History of Chinese Classical Scholarship, Volume II: Canon and Commentary*; and the co-authored monograph *New Ideas Concerning Arts and Social Studies*. She has also published multiple papers in peer-reviewed international academic journals.

Dr. Huang is currently a core member of the Major Project of the National Social Science Fund of China (Project No. 23&ZD236), titled "A Study of the Hermeneutical History of Chinese Classical Scholarship and the Collation of Exegetical Texts." The interim result of this project—*The Triple Interaction of Classics, History, and Literature: A New Interpretation of Li Zhi's Four Books Critique from Global Perspectives*—is forthcoming.

Dr. Huang's teaching, research, and curatorial practice continuously bridge Eastern and Western critical traditions, exploring the intersections of art, gender, ecology, hermeneutics and language education studies, within an increasingly globalized academic landscape. Her curated exhibitions, including *Pink Bomb*, *The Vivid Hues from the Deeper Depths of Dreams*, *A Forest on Paper*, *Dreams Within Pages*, and *Phantasmagoric Visions*, are highly experimental and pioneering in nature.

*In accordance with Chinese naming conventions, the author's family name is placed before her given name as a sign of respect for her cultural tradition.

Chapter 1
Introduction

Huang Xiao

Abstract: This study investigates the integration of generative AI within a junior high Chinese as a Second Language (CSL) classroom to address common psychological barriers like high student anxiety. By implementing an 18-week curriculum featuring a deliberately "mistake-prone" AI partner, the research explores how intentional errors trigger negotiation of meaning and language output modification. The study further tracks the evolution of student role perceptions—shifting from viewing AI as a "silly robot" to a customizable tool—while documenting essential ethical and privacy protection strategies. Findings aim to contribute to Second Language Acquisition theories by providing preliminary evidence for the applicability of the Interaction and Output Hypotheses in human-machine environments. Ultimately, the research provides CSL practitioners with actionable teaching strategies and a methodological template for studying technology-rich, heterogeneous classrooms.

1 Research Background

1.1 A New Era of AI Technology in Language Education

Since the end of 2022, generative artificial intelligence (AI) technologies, represented by ChatGPT, have entered the field of education at an unprecedented pace. Unlike previous educational technologies, generative AI possesses capabilities such as natural language conversation, real-time feedback, content generation, and role-playing, which endow it with unique potential in language teaching. In contrast to traditional language learning software (e.g., Duolingo, Rosetta Stone), which primarily provides preset exercises and standardized feedback, generative AI can en-

gage students in open-ended, contextualized dialogues and can even play specific roles based on instructional needs (e.g., a restaurant server, a passerby asking for directions, a demanding customer).

This technical characteristic aligns closely with core concerns in the field of second language acquisition (SLA). For a long time, creating ample, low-anxiety, and meaningful opportunities for language output has been a persistent challenge in teaching practice. In traditional classrooms, it is difficult for teachers to engage in one-on-one dialogue practice with all students simultaneously. While peer-to-peer dialogue has value, it often fails to provide effective correction and input due to similar proficiency levels between students. The advent of AI, at least technologically, offers new possibilities to address this dilemma.

However, the integration of technology into classrooms is never a simple process of "tool substitution." How should AI be used in language classrooms? How do students perceive and respond to this "non-human" conversation partner? What privacy and ethical issues arise from using AI? These questions urgently need to be answered through detailed, context-rich qualitative research conducted in real classrooms, especially given the current pace of technological iteration.

1.2 The Specific Challenges Facing Junior High Second Language Learners

This study focuses on junior high school learners of Chinese as a second language (CSL). This age group (12–15 years) faces unique psychological and social challenges in language learning.

First, junior high students are in a period of rapid self-awareness development. They are highly sensitive to peer evaluation and fear making mistakes in public. In the language classroom, this means that "speaking up" itself becomes a significant psychological barrier. Many junior high CSL learners know the correct expressions but are afraid to utter them in front of classmates and the teacher—they worry about being laughed at for non-native pronunciation, being corrected for grammatical errors, and "losing face." This anxiety is described in Krashen's (1982) Affective Filter Hypothesis as an "affective filter" that is too high, hindering the transformation of language input into acquisition.

Second, while junior high students' cognitive abilities have developed enough to handle abstract grammatical rules and complex communicative tasks, their attention span and self-control remain limited. Traditional grammar-translation methods and mechanical drills are prone to bore them, thereby reducing learning motivation. Balancing "meaningfulness" and "enjoyability" is a core challenge in junior high second language teaching.

Third, in international school settings, CSL students often come from diverse linguistic and cultural backgrounds. Their native languages may include English, Korean, Japanese, French, etc., and their length of Chinese study ranges from a few months to several years. This heterogeneity makes it difficult for teachers to meet

all students' needs in a single lesson—a task that is too easy for some students may be too difficult for others.

These challenges collectively point to a need: junior high CSL classrooms require a tool that can provide ample, personalized, low-anxiety, and real-time oral practice opportunities. AI, at least theoretically, has the potential to meet this need.

1.3 The Classroom Context and Curriculum Design of This Study

The field site for this study is a junior high CSL classroom at an international school. Students are aged 12 to 14, with Chinese proficiency approximately at HSK Level 3 (vocabulary of about 600 words, capable of simple daily communication). The class size is 14 students, representing seven different native language backgrounds.

During the fall semester of the 2025 academic year, I designed and implemented an 18-week AI-assisted Chinese curriculum. The core design principles of the curriculum included:

- **AI as a "Mistake-Prone Language Partner"**: Contrary to the traditional pursuit of a "perfect AI," the AI in this study was intentionally set to be a "silly" character—it makes predictable errors, "fails to understand" certain student expressions, and even provides incorrect directions or information. This design aims to actively create opportunities for negotiation of meaning, transforming students from the "corrected" to the "correctors."
- **Human-AI Collaborative Task Structure**: Each AI-integrated lesson followed a five-stage structure: Warm-up (AI pronunciation assessment) – Presentation + AI demonstration – Pair-based AI dialogue – AI game for consolidation – AI-assisted output. AI was not the entirety of the lesson but was interwoven as one component among real-person interaction, writing exercises, and cultural activities.
- **Progressive Adjustment of AI Roles**: Over the semester, the AI's role transitioned from a "silly robot" (weeks 1–6) to a "normal assistant" (weeks 7–12) and finally to a "student-customizable tool" (weeks 13–18). This design aimed to observe the trajectory of changes in students' perceptions of the AI's role.
- **Embedded Privacy and Ethical Design**: All AI activities did not use students' real names, did not save conversation logs, and all data were cleared after each class. Parents received a detailed explanatory letter at the beginning of the semester and had the right to opt out.

This curriculum design provided rich data sources for this study—18 weeks of real classroom observations, student-AI dialogue logs, student task cards, and interviews with students and parents.

2 Research Purpose and Research Questions

2.1 Research Purpose

This study aims to develop an in-depth understanding of the actual processes of AI-assisted language teaching in junior high CSL classrooms, focusing on three interrelated dimensions:

First, the interactional dimension: How do students negotiate meaning with the AI? When the AI "fails to understand" or "makes a mistake," how do students respond? Do these interactions promote students' language output and modification?

Second, the perceptual dimension: How do students perceive the AI's role in their Chinese learning? Is it a tool, a peer, an opponent, or something else? How does this perception evolve over time and with interactive experience?

Third, the ethical dimension: This dimension is further divided into two sub-questions to reflect its dual focus—teacher practices and parent perspectives.

> **RQ3a**: What privacy and ethical protection strategies does the teacher employ in AI-assisted teaching? How does the teacher balance the technological potential of AI with ethical concerns?
>
> **RQ3b**: What narratives of concern and acceptance do parents exhibit regarding the introduction of AI into the Chinese classroom?

These two sub-questions are analyzed together in Chapter 6 because they are empirically interlinked: parents' concerns shaped teacher practices, and teacher responses, in turn, influenced parents' acceptance. However, they are distinguished analytically to maintain clarity between the teacher's actions (RQ3a) and parents' perspectives (RQ3b).

Through detailed description and analysis of these three dimensions, this study aims to provide evidence-based, theoretically grounded insights for the integration of AI into language classrooms, rather than a simplistic "effectiveness evaluation" or "tool recommendation."

2.2 Research Questions

Based on the above purpose, this study proposes the following three core research questions:

RQ1 (Interactional dimension): In an AI-assisted junior high CSL classroom, what kinds of meaning negotiation behaviors occur between students and the AI? How does the AI's "intentional error-making" setting affect students' language output and modification?

RQ2 (Perceptual dimension): How do students perceive the AI's role in their Chinese learning? How does this role perception evolve over time (one semester) and across interactive situations?

RQ3 (Ethical dimension): What privacy and ethical protection strategies does the teacher employ in AI-assisted teaching? What narratives of concern and acceptance do parents exhibit regarding the introduction of AI into the Chinese classroom?

These three research questions are all addressed within a qualitative research paradigm. This study does not pursue statistical "representativeness" or "significance" but instead aims, through in-depth, context-rich description, to reveal the complexity and richness of AI-assisted language teaching.

3 Significance of the Study

3.1 Theoretical Significance

This study potentially contributes to theories of second language acquisition and computer-assisted language learning in the following respects:

First, it extends the boundaries of the Interaction Hypothesis. Long's (1996) Interaction Hypothesis is primarily based on empirical studies of human-human interaction. Do the characteristics of "human-machine meaning negotiation" differ from those of "human-human meaning negotiation"? Does the AI's "predictable error" trigger students' output modification more effectively than the "natural errors" in real-person conversations? These questions remain underexplored. Through detailed descriptions of human-machine interaction in real classrooms, this study provides preliminary evidence for the applicability of the Interaction Hypothesis in the AI era.

Second, it enriches frameworks of AI roles. Luckin et al. (2016) proposed four roles for AI in education (tutor, coach, peer, tool), but this framework is primarily based on theoretical reasoning and lacks qualitative validation in real classrooms. Through long-term classroom observations, this study describes how students dynamically assign roles to the AI in actual interactions, proposing the concept of "role fluidity" to supplement and revise this theoretical framework.

Third, it connects the Output Hypothesis with AI pedagogical practice. Swain's (2005) Output Hypothesis emphasizes three functions: "noticing trigger," "hypothesis testing," and "metalinguistic reflection." By analyzing students' output modification behaviors in interaction with the AI, this study explores whether and how the AI triggers these processes, providing empirical cases for research on the Output Hypothesis in human-machine interaction environments.

3.2 Practical Significance

The practical significance of this study is mainly reflected in the following aspects:

First, it provides international school CSL teachers with actionable AI teaching strategies. The 18-week curriculum design, detailed lesson plans, task card templates, parent letters, etc., from this study can be adapted and used by other teachers. In particular, the design concept of "AI as a mistake-prone language partner" contrasts with the traditional pursuit of a "perfect AI" and may offer new pedagogical inspiration to other educators.

Second, it provides user-perspective feedback for the design and improvement of AI educational products. This study describes how students actually use (and do not use) the AI, under what circumstances the AI is helpful, and under what circumstances it becomes an obstacle. These findings can inform the design of AI products for K-12 educational contexts.

Third, it provides real-classroom-based ethical considerations for schools developing AI usage policies. The parent letters, privacy protection measures, and concerns expressed in parent interviews presented in this study can serve as references for other schools when formulating similar policies.

3.3 Methodological Significance

This study adopts a purely qualitative research paradigm, using multiple data sources including classroom observations, semi-structured interviews, student task analyses, and researcher reflective journals. By demonstrating how to progressively "fill in" observation record sheets over 18 weeks, how to extract themes from narrative descriptions, and how to manage the researcher's dual role as both teacher and researcher, this study aims to provide a referential case for novice researchers engaged in qualitative educational research.

Besides, it provides a replicable template for qualitative researchers studying technology-rich classrooms. The multi-source data collection strategy, the teacher-researcher reflexivity protocols, and the thematic analysis procedures detailed in Chapter 3 can serve as a model for other classroom-based qualitative studies. One particularly transferable methodological feature is the use of the "teacher reflective journal" (see Appendix G). Unlike a general research journal, this structured template separates lesson overview, technical operations, student observations, teaching decisions, ethical dilemmas, and reflexivity—allowing the teacher-researcher to document both descriptive events and interpretive reflections in parallel. Novice qualitative researchers can adopt this template and adapt it to their own teaching-research contexts. The transparency with which the researcher's dual role is managed—including explicit discussion of tensions and mitigation strategies—addresses a gap in the qualitative technology-and-learning literature, where researcher positionality is often under-acknowledged.

4 Definition of Core Concepts

To ensure clarity and dialogic engagement, this study defines the following core concepts:

AI-assisted language teaching	Refers to the use of AI tools (in this study, the voice version of ChatGPT) as pedagogical aids in the language teaching process, providing functions such as dialogue practice, real-time feedback, and content generation. The AI use in this study is not "AI-led teaching" but "teacher-designed, pedagogically targeted, short-duration AI intervention."
Negotiation of meaning	Borrowed from Long's (1996) definition, refers to the interactional adjustments made by learners to overcome comprehension difficulties during communication, including clarification requests, confirmation checks, repetitions, paraphrasing, and simplification. In this study, negotiation of meaning specifically refers to interactional adjustments occurring between students and the AI due to comprehension obstacles.
Output modification	Borrowed from Swain's (2005) concept, refers to learners' adjustments to their output when noticing a discrepancy between their language production and target language forms. In this study, output modification is operationalized as students altering their original expression in response to the AI's "failure to understand" or "erroneous response."
AI role perception	Refers to the "identity" or "role" that students assign to the AI during interaction, such as tool, peer, opponent, etc. This concept emphasizes that the AI's role is not technologically predetermined but is dynamically constructed by students in interaction.
Affective filter	Borrowed from Krashen's (1982) concept, refers to the effect of affective factors such as anxiety, self-confidence, and motivation on the efficiency of language input absorption. In this study, the affective filter is operationalized as students' self-reported or observer-inferred level of anxiety/relaxation during dialogues with the AI.

AI literacy	Refers to students' ability to use AI tools effectively, critically, and responsibly. In this study, AI literacy is operationalized as students' ability to give instructions to the AI, correct AI errors, evaluate AI outputs, and protect personal privacy when interacting with the AI.
Teacher-researcher dual role	In this study, the researcher is simultaneously the class's Chinese teacher and curriculum designer. This identity has advantages (deep access to the field, long-term observation, trust-building) but also limitations (potential bias, students potentially changing behavior due to the researcher's presence). Chapter 3 discusses strategies for managing this dual role in detail.

5 Book Structure

This book comprises seven chapters:

- **Chapter 1: Introduction**: Presents the research background, research purpose and questions, significance of the study, and definitions of core concepts.
- **Chapter 2: Literature Review**: Systematically reviews second language acquisition theories (Interaction Hypothesis, Output Hypothesis, Affective Filter Hypothesis), computer-assisted language learning and AI role frameworks, and privacy and ethics research in educational AI, identifying gaps in existing research and the theoretical positioning of this study.
- **Chapter 3: Research Methodology**: Details the research design, participants, data collection methods, data analysis methods (thematic analysis), the researcher's role and reflexivity, and ethical considerations.
- **Chapter 4: Research Findings I – Human-Machine Negotiation of Meaning**: Presents qualitative descriptions of meaning negotiation behaviors between students and the AI, including types and modes of negotiation, and successful and unsuccessful cases.
- **Chapter 5: Research Findings II – The Fluidity and Evolution of AI Roles**: Presents students' perceptions of the AI's role and their trajectories of change over time.
- **Chapter 6: Research Findings III – Privacy, Ethics, and Digital Citizenship Education**: Presents the teacher's ethical practices and narratives of parental concern.
- **Chapter 7: Discussion and Conclusion**: Engages in dialogue with theory, summarizes research findings, proposes pedagogical recommendations, and identifies research limitations and future directions.

6 Reader Navigation: How This Book Answers Its Research Questions

The table below shows where each research question is addressed and which data sources support it.

Table 1 Reader Navigation

Research Question	Addressed	Primary Data Sources
RQ1: Meaning negotiation behaviors and output modification	Ch. 4	Classroom observations, dialogue logs, task cards
RQ2: AI role perception and its evolution	Ch. 5	Student interviews, observations, task card comments
RQ3: Teacher ethical practices and parent narratives	Ch. 6	Teacher reflective journals, parent interviews, parent questionnaires

Readers interested in practical classroom materials may also consult Appendices A–I, which contain parent letters, task card templates, interview protocols, observation forms, and coding schemes.

7 Conceptual Framework: How Research Questions, Theories, and Findings Connect

The section below illustrates the relationship between this study's three research questions, the theoretical frameworks that inform them, and the chapters where findings are presented. This framework uses a nested structure to show how the three theoretical streams overlap and inform one another, with their intersection representing the integrated analytical lens applied in Chapter 7.

7.1 Theoretical Frameworks and Their Corresponding Research Questions

1. **Interaction Hypothesis** (Long, 1996)
   → Informs RQ1: Meaning negotiation
   → Key concepts: negotiation of meaning, clarification requests, confirmation checks
2. **Output Hypothesis** (Swain, 2005)
   → Informs RQ1: Output modification
   → Key concepts: noticing trigger, hypothesis testing, metalinguistic reflection

3. **Affective Filter Hypothesis** (Krashen, 1982)
   → Informs RQ1: Affective dimensions of interaction
   → Key concepts: low-anxiety environment, motivation, self-confidence
4. **AI Role Framework** (Luckin et al., 2016)
   → Informs RQ2: AI role perception
   → Key concepts: tutor, coach, peer, tool
5. **RRI Framework** (von Schomberg, 2013) **+ OECD Privacy Principles** (2022)
   → Informs RQ3: Ethics and privacy
   → Key concepts: informed consent, data protection, parental engagement, reflexivity

How the Theories Overlap (The Central Intersection)

All five theoretical streams converge on the core phenomenon of this study: AI-mediated CSL learning in a real junior high classroom.

At the intersection of these theories, this study investigates:

- **RQ1** (Chapter 4): How students negotiate meaning with a deliberately "imperfect" AI and modify their output in response to AI feedback.
- **RQ2** (Chapter 5): How students perceive the AI's role (tool, peer, opponent, etc.) and how these perceptions shift across time and interactional contexts.
- **RQ3** (Chapter 6): What ethical practices teachers employ and what narratives of concern and acceptance parents express.

7.2 From Theories to Findings to Theoretical Contributions

Chapter 4 (Findings I): Human-Machine Negotiation of Meaning
⌞ Draws on: Interaction Hypothesis + Output Hypothesis + Affective Filter

Chapter 5 (Findings II): Role Fluidity and Evolution of AI Roles
⌞ Draws on: AI Role Framework (Luckin et al., 2016)

Chapter 6 (Findings III): Privacy, Ethics, and Parental Communication
⌞ Draws on: RRI Framework + OECD Privacy Principles

Chapter 7 (Discussion and Conclusion)
⌞ Integrates all three findings streams to propose:

1. Human-machine negotiation of meaning — An extension of Long's (1996) Interaction Hypothesis to AI contexts
2. Role fluidity — A dynamic supplement to Luckin et al.'s (2016) AI role framework
3. Opponent-to-partner evolution trajectory — A new model tracing lower-proficiency learners' changing perceptions of AI (Testing Machine → Fallible Opponent → Game Partner → Learning Partner)

7.3 Summary Table: Research Questions, Theories, and Chapters

Table 2 Research Questions, Theories, and Chapters

Research Question	Primary Theoretical Framework	Chapter
RQ1 (Negotiation & Output)	Interaction Hypothesis (Long, 1996); Output Hypothesis (Swain, 2005); Affective Filter (Krashen, 1982)	Ch. 4
RQ2 (Role Perception)	AI Role Framework (Luckin, Holmes, Griffiths, & Forcier, 2016)	Ch. 5
RQ3 (Ethics & Privacy)	RRI Framework (von Schomberg, 2013) + OECD Privacy Principles (2022)	Ch. 6

All three streams converge in Chapter 7 for discussion, theoretical integration, and pedagogical recommendations.

8 Chapter Summary

This chapter begins with the background of AI technology entering a new era in language education, points out the specific challenges faced by junior high CSL learners (high anxiety, susceptibility to boredom, heterogeneity), introduces the classroom context of this study and the basic framework of the 18-week AI-assisted curriculum. On this basis, it proposes three core research questions (interaction, perception, ethics), elaborates on the theoretical, practical, and methodological significance of the study, defines key concepts, and previews the structure of the thesis.

This study is not a simple answer to the question "Is AI effective?" but rather an in-depth description and interpretation of "how AI is used, perceived, and managed in real classrooms." The central concern of the research is: in an era of rapid technological change, how can we responsibly and wisely integrate AI into language education so that it truly serves learners' development, rather than becoming yet another "decorative technological frill"?

Chapter 2
Literature Review

Huang Xiao

Abstract: This chapter systematically reviews the theoretical lineages and empirical research findings relevant to the three core research questions of this study.

1 Introduction

This chapter is divided into four main sections.

The first section reviews three classic theories from the field of second language acquisition—Long's (1996) Interaction Hypothesis, Swain's (1985, 2005) Output Hypothesis, and Krashen's (1982) Affective Filter Hypothesis. These three theories constitute the theoretical foundation for understanding the process of "human-machine negotiation of meaning" in this study.

The second section traces the developmental trajectory of the field of computer-assisted language learning (CALL), focusing on Luckin et al.'s (2016) framework of AI roles and subsequent extensions by later researchers, providing conceptual tools for this study's exploration of "students' perceptions of AI roles."

The third section reviews research on privacy and ethics in educational AI, focusing on ethical principles for AI use in K-12 educational contexts, parental concerns, and the Responsible Research and Innovation (RRI) framework, providing theoretical support for this study's exploration of "teachers' ethical practices and parent narratives."

Huang Xiao • Krirk University, Bangkok, Thailand
e-mail: huang.xiao@krirk.ac.th

© The Author(s). Published in the USA.
https://doi.org/10.63408/979-8-9928364-4-8

The fourth section identifies gaps in existing research and articulates the theoretical positioning of this study, clarifying its place in theoretical dialogue and its potential contributions.

2 Second Language Acquisition Theories: Interaction, Output, and Affect

2.1 The Interaction Hypothesis: Negotiation of Meaning and Language Acquisition

The Interaction Hypothesis, proposed by Michael Long in the 1980s and systematically elaborated in 1996, holds that second language acquisition occurs when learners engage in "negotiation of meaning." Recent research has extended this hypothesis to human-AI interaction contexts. Chen et al. (2024) found that learners who use ChatGPT as a conversational partner engage in meaning negotiation behaviors—including clarification requests and confirmation checks—that parallel those observed in human-human interaction, while also benefiting from the AI's infinite patience and non-judgmental feedback. O'Connor (2025) further argues that AI systems can function as "fellow participants" in the language learning process, creating unique opportunities for negotiation that differ from traditional human interlocutor settings.

Long (1996, p. 418) defines negotiation of meaning as "interactional adjustments made by learners to overcome communication difficulties," including clarification requests (e.g., "What did you say?"), confirmation checks (e.g., "Do you mean…?"), and repetition (e.g., "Turn left? Turn left?"). When communication breaks down, learners are "pushed" to adjust their language output so that the interlocutor can understand. This process is considered a key mechanism promoting acquisition.

The theoretical roots of the Interaction Hypothesis can be traced to Krashen's Input Hypothesis. Krashen (1982) argued that the only condition for acquisition is that learners are exposed to "comprehensible input," i.e., input slightly above their current level (i+1). However, Long (1996) further argued that providing comprehensible input alone is insufficient—learners need to "negotiate" the meaning of input through interaction for it to be truly understood and internalized.

Empirical research on the Interaction Hypothesis has primarily been based on human-human interaction in classroom and laboratory settings. For example, Pica (1994) found that when learners have opportunities to negotiate meaning with conversation partners, they are more likely to notice their language errors and correct them. Gass and Varonis (1994) showed that the frequency and quality of meaning negotiation are positively correlated with learners' language progress.

However, the research tradition of the Interaction Hypothesis has focused mainly on human-human interaction (e.g., native speaker–non-native speaker or non-native

speaker–non-native speaker). When the conversation partner is not human but an AI, does meaning negotiation exhibit the same characteristics? Does the AI's "predictable error" trigger students' output modification more effectively than the "natural errors" in real-person conversations? These questions have not been adequately addressed in existing literature. This study aims to respond to this question through observations of "human-machine negotiation of meaning" in real classrooms.

2.2 The Output Hypothesis: Noticing Trigger, Hypothesis Testing, and Metalinguistic Reflection

The Output Hypothesis, proposed by Merrill Swain in 1985 and systematically updated in 2005, began as a critique of Krashen's Input Hypothesis for overemphasizing the role of input while neglecting the value of output.

Swain (2005, p. 474) argues that output serves at least three independent functions in second language acquisition:

First, the noticing/triggering function: When learners attempt to produce output in the target language, they notice the gap between what they "want to say" and what they "can say." This noticing is a prerequisite for acquisition—learners must first notice a problem before they can attempt to solve it.

Second, the hypothesis-testing function: Learners use output as a means to test their hypotheses about the target language. For example, a learner uncertain about a grammatical form can speak it and observe the conversation partner's reaction to test the hypothesis. If the partner shows non-understanding or provides correction, the learner adjusts the hypothesis.

Third, the metalinguistic reflective function: When learners are asked to explain or reflect on their output, they use metalinguistic knowledge to analyze and solve problems. This process helps transform explicit knowledge into implicit knowledge.

Applications of the Output Hypothesis in classroom teaching have been extensively studied. For instance, Swain and Lapkin (1995) found that when students were asked to collaboratively complete writing tasks, the "language-related episodes" they engaged in during output—interactions involving discussion and modification of language forms—were positively correlated with language acquisition.

This study pays particular attention to the manifestation of the Output Hypothesis's "noticing trigger" and "hypothesis-testing" functions in human-machine interaction. When the AI says "I don't understand" or gives an erroneous response, do students "notice" the gap between their output and the AI's expectation? Do they test their language hypotheses by modifying their output? These questions will be addressed through classroom observations and student interviews.

2.3 The Affective Filter Hypothesis: Anxiety, Motivation, and Language Acquisition

The Affective Filter Hypothesis is one of the five hypotheses of Krashen's (1982) Monitor Model. This hypothesis argues that learners' affective states—including anxiety, motivation, and self-confidence—affect the efficiency with which language input is converted into acquisition.

Krashen (1982, p. 31) metaphorically describes the "affective filter" as an invisible barrier. When learners have high anxiety, low motivation, and low self-confidence, the affective filter is "high," blocking input from reaching the language acquisition device. Conversely, when learners have low anxiety, high motivation, and high self-confidence, the affective filter is "low," allowing input to be absorbed smoothly.

The Affective Filter Hypothesis has had a profound impact on language teaching practice. It reminds teachers that creating a low-anxiety classroom environment is as important as providing high-quality language input. In traditional language classrooms, student anxiety often stems from fear of "being corrected"— speaking incorrectly in public invites correction from the teacher and ridicule from peers. This social anxiety is a major source of the affective filter.

In recent years, researchers have begun exploring how technology might lower the affective filter in language learning. For example, Godwin-Jones (2022) notes that AI conversation partners have a "non-judgmental" quality—AI does not show impatience, does not correct publicly, and has no facial expressions— which may make learners feel safer and more willing to take risks. However, these discussions remain largely theoretical, lacking qualitative evidence from real classrooms. This study will test this hypothesis through students' affective responses to the AI.

2.4 Implications of the Theoretical Review for This Study

The three theories reviewed above provide conceptual tools for understanding the process of "human-machine negotiation of meaning" in this study:

The Interaction Hypothesis directs attention to what kinds of meaning negotiation behaviors occur between students and the AI, and how these behaviors affect students' language adjustments.

The Output Hypothesis provides a specific framework for analyzing "output modification"—noticing trigger, hypothesis testing, metalinguistic reflection—three concepts that can be operationalized as observable behaviors in classroom observations.

The Affective Filter Hypothesis guides attention to students' affective experiences, i.e., whether they feel more relaxed or more anxious when conversing with AI, and how this affective state affects their engagement and learning behaviors.

However, existing theories are primarily based on research on human-human interaction. Whether they are applicable to human-machine interaction contexts is an empirical question. This study does not simply "apply" these theories to the AI context but aims to engage in dialogue with them through real classroom observations, exploring their boundaries of applicability and possible extensions.

3 Computer-Assisted Language Learning and AI Roles

3.1 The Trajectory of CALL: From Behaviorism to Generative AI

Computer-Assisted Language Learning (CALL) as a research field has a history of over half a century. Warschauer and Healey (1998) proposed a widely cited three-stage model describing the theoretical evolution of CALL:

First stage: Behavioristic CALL (1960s–1970s). CALL software was based on behaviorist learning theory, emphasizing stimulus-response-reinforcement. Typical applications were "drills and practice" software—the computer presented a stimulus (e.g., a translation of an English word into Chinese), the student entered an answer, and the computer gave "correct" or "incorrect" feedback. This model was theoretically grounded in Skinner's operant conditioning.

Second stage: Communicative CALL (1980s–1990s). With the rise of communicative language teaching, CALL software began to emphasize language use in authentic contexts. Software in this stage was no longer limited to grammar drills but provided more open-ended tasks (e.g., simulated dialogues, role-plays). However, Warschauer and Healey (1998) noted that much software marketed as "communicative" remained behavioristic in essence—it merely packaged drills in more attractive interfaces.

Third stage: Integrative CALL (2000s–2010s). This stage is characterized by technology being "integrated" into all aspects of language teaching, rather than existing as separate "computer classes." Tools such as the internet, multimedia, and social media were used to support authentic language use and intercultural communication.

Warschauer and Healey's (1998) three-stage model is valuable for understanding the historical evolution of CALL, but it was proposed before the advent of generative AI.

After 2022, generative AI represented by ChatGPT entered language classrooms. These tools are capable not only of providing prescriptive feedback but also of engaging in open-ended, contextualized dialogue. This suggests that CALL may have entered a fourth stage—Generative CALL. Recent scholarship has begun to map this new terrain. Fryer et al. (2020) examined student interactions with chatbots, finding that learners' perceptions of the machine as a "human-like" partner significantly affected engagement. Kim et al. (2022) investigated AI-powered feedback tools, noting that the immediacy and consistency of AI feedback can complement—but

not replace—teacher feedback. Burston (2023) provided a critical overview of AI in language learning, cautioning against technological determinism. Most relevant to this study, Huang et al. (2023) and Yan (2023) explored the pedagogical dimensions of large language models in language classrooms, calling for more qualitative, classroom-based research on how learners actually interact with generative AI. The present study responds directly to that call through its 18-week longitudinal design in a real international school classroom.

In this stage, AI is no longer a tool executing preset programs but a dynamic conversation partner capable of "generating" new language content, playing specific roles, and even "making mistakes." This study is situated within this new technological context.

3.2 Four Roles of AI in Education

Luckin et al. (2016), commissioned by the UK Open University and Pearson, authored a report titled *Intelligence Unleashed: An Argument for AI in Education*. The report proposed four roles for AI in education:

Role 1: Tutor. AI as tutor assesses the learner's current knowledge state, provides personalized instructional content and feedback, and guides the learner along an optimal learning path. AI technologies corresponding to this role include Intelligent Tutoring Systems and adaptive learning platforms. The characteristic feature of the tutor role is that the AI knows the answers, and the learner is the "one being taught."

Role 2: Coach. AI as coach observes the learner's learning process and performance, providing hints, strategic suggestions, and encouragement without excessive intervention. The distinction between coach and tutor is that the coach does not give answers directly but helps learners find answers themselves. This role is more common in sports training and vocational skills development.

Role 3: Peer. AI as peer collaborates equally with learners to complete tasks, potentially sharing its own "thoughts" and "confusions," and may even make mistakes. The core characteristic of the peer role is that the relationship between learner and AI is collaborative and non-hierarchical. The AI is not a "smarter tutor" but a "partner to learn with."

Role 4: Tool. AI as tool is actively manipulated and used by learners to complete specific tasks (e.g., translation, word lookup, text generation). In this role, the learner has complete agency, and the AI is a passive executor.

Luckin et al.'s (2016) framework provides clear conceptual tools for understanding the functions of AI in education. However, the framework is primarily based on theoretical reasoning and functional analyses of existing AI technologies, lacking qualitative research on "how students actually perceive AI roles" in real classroom environments. More importantly, the framework tends to treat AI roles as fixed—a given AI system is either a tutor or a peer. However, this study hypothesizes that the same AI tool in the same lesson may play different roles in different students' eyes,

and may even undergo "role fluidity" in the same student's eyes as interactional contexts shift. This hypothesis requires empirical data to test.

3.3 Extensions and Debates on the AI Role Framework

Following the proposal of Luckin et al.'s (2016) framework, subsequent researchers have critiqued and extended it.

First, Kukulska-Hulme and Shield (2008), in the context of mobile-assisted language learning (MALL), argued that the role of technology is often not predetermined by designers but "negotiated" by learners through use. This perspective is related to the concept of "user agency"—learners are not passive recipients of technologically assigned roles but actively construct their relationship with technology. This study adopts this perspective, treating AI role perception as an interactional achievement rather than a technological given.

Second, Godwin-Jones (2022), discussing generative AI and language learning, noted that AI roles may be more complex than Luckin et al.'s (2016) framework suggests. For example, when AI simultaneously provides feedback (tutor function), encourages the learner (coach function), and participates in collaborative dialogue (peer function), its role is "hybrid" rather than singular. Godwin-Jones (2022, p. 15) observes that the relationship between learners and AI is not one of simple either/or categories: learners may, in the same conversation, treat AI as a tool, a peer, or even as something to be educated.

Third, Tsai (2021), in a study on chatbot-assisted English learning, proposed the concept of "AI as a mistake-prone conversation partner." Tsai found that when chatbots were set to be "perfect" (i.e., always giving correct responses), learner engagement was lower; when chatbots were set to "make mistakes," learners were more willing to actively correct them, and engagement was higher. Tsai (2021, p. 1088) argues that the AI's apparent weakness becomes a pedagogical advantage, giving learners an opportunity to become "experts."

Tsai's (2021) study is highly consistent with the curriculum design philosophy of this study—the AI in this study was intentionally set as a "silly" character precisely to create opportunities for learners to become "correctors." However, Tsai's (2021) study was a short-term laboratory study (approximately 15 minutes per session), whereas this study is a long-term observation in real classrooms (one semester, 18 weeks). This allows this study to track the evolution of AI role perception, rather than merely describing a snapshot.

3.4 Implications of the Theoretical Review for This Study

Luckin et al.'s (2016) AI role framework and extensions by subsequent researchers provide conceptual tools for analyzing "students' perceptions of AI roles":

The four role types (tutor, coach, peer, tool) provide a starting point for preliminary classification of student perceptions.

The concept of role fluidity (Godwin-Jones, 2022) directs attention to the dynamic nature of role perception.

The concept of AI as a mistake-prone peer (Tsai, 2021) is directly relevant to this study's curriculum design, providing an empirical foundation for dialogue.

However, existing research is primarily based on short-term experiments or theoretical reasoning, lacking long-term tracking of AI role perception in real classroom environments. This study will describe the trajectory of changes in students' AI role perceptions through 18 weeks of classroom observations and three interviews (week 6, week 12, week 18), thereby supplementing this theoretical framework. In Chapter 5, this study will propose the concept of "role fluidity" as an extension to Luckin et al.'s framework. Role fluidity refers to the phenomenon whereby the same student, in the same lesson, may assign different roles to the AI depending on changes in the interactional context.

4 Privacy and Ethics in Educational AI

4.1 The Responsible Research and Innovation Framework

Responsible Research and Innovation (RRI) is a policy framework proposed by the European Union in the 2010s to ensure that technological innovation considers ethical, social acceptability, and public interest at every stage of development. Von Schomberg (2013, p. 63) defines RRI as "a transparent, interactive process by which societal actors and innovators respond to each other to ensure that the innovation process and its products are ethically acceptable and socially desirable."

The RRI framework typically includes five core dimensions (Stilgoe et al., 2013):

- **Anticipation**: Foreseeing potential social, ethical, and environmental consequences before innovation occurs.
- **Reflexivity**: Innovators reflecting on their own assumptions, values, and potential biases.
- **Inclusion**: Involving diverse stakeholders (including end-users) in the innovation process.
- **Responsiveness**: Adjusting the direction and strategy of innovation based on new evidence and public concerns.
- **Sustainability**: Considering the long-term social impacts of innovation.

In the field of educational AI, the RRI framework has been used to analyze and guide the development of AI educational products. For example, Holmes et al. (2022), in a report titled *Ethics of AI in Education: Towards a Community-Wide Framework*, used the RRI framework to identify six major ethical challenges facing educational AI. More recently, Penuel et al. (2025) have extended this line of

inquiry by proposing a framework that attends not only to "how" AI is designed but also to "for what," "for whom," and "with whom"—emphasizing the importance of involving diverse stakeholders (including students, teachers, and parents) in the design process and aligning AI with ethical principles that support democratic and collaborative learning. This stakeholder-first approach resonates directly with the methodological choices made in this study, including the use of parent letters, informed consent, and ongoing communication with families throughout the 18-week intervention.

This study uses the RRI framework as a theoretical tool for analyzing teachers' ethical practices. Specifically, the parent letter, privacy protection measures (no logging, no conversation saving, use of pseudonyms), and "opt-out" option in this study can be understood as operationalizations of the RRI framework's dimensions of "inclusion," "responsiveness," and "reflexivity."

Beyond the RRI framework, two additional ethical frameworks are relevant to this study. First, UNESCO's (2021) *Recommendation on the Ethics of Artificial Intelligence* provides a global policy framework emphasizing human rights, inclusion, and gender equality. It calls for "AI literacy for all" and for educational AI to be "human-centered"—principles directly reflected in this study's parent communication and student privacy protections. Second, Selwyn (2019) has critically examined the "datafication" of education, warning that educational technology often prioritizes data extraction over student welfare. His work reminds us that ethical AI use is not simply about following checklists (e.g., "clear dialogue logs") but about maintaining a critical stance toward technology's role in education. This study's teacher reflective journals—which document ongoing ethical reflection rather than one-time compliance—embody this critical stance.

4.2 Privacy Principles for Educational AI

AI use in educational contexts involves large amounts of sensitive data— students' voices, writing, learning progress, and even emotional states. The OECD (2022) released a report titled *Privacy Framework for AI in Education*, proposing the following core principles:

- **Minimal Collection**: Educational institutions should collect only the minimum data necessary to achieve educational objectives. For example, if student identification is not needed, students should not be required to provide their names.
- **Transparency**: Educational institutions should clearly inform students and parents about what data is collected, how it is used, how long it is stored, and with whom it is shared.
- **Data Deletion**: Educational institutions should establish clear data retention periods and securely delete data after those periods expire. Students or parents should also have the right to request deletion of personal data.

- **Purpose Limitation**: Collected data may only be used for the stated educational purposes and not for other uses (e.g., behavioral analysis, commercial marketing).
- **Parental Consent**: For K-12 students (especially younger students), informed consent from parents must be obtained before using any AI tool that collects personal data.
- **Data Security**: Educational institutions should take technical and administrative measures to prevent data breaches, misuse, or unauthorized access.

The design of this study—not using students' real names, not saving conversation logs, clearing data after class, and providing a parent letter with clear information and consent—follows the above principles. However, implementing these principles in real classrooms is not without challenges. For example, the "minimal collection" principle may be in tension with "the need to collect data for classroom observation research." Chapter 3 (Research Methodology) discusses in detail how this tension is addressed.

4.3 Parental Concerns About Educational AI

Parental attitudes toward the use of AI technology in schools is an emerging research area. Akgun and Greenhow's (2022) review of parental concerns found that worries cluster mainly around the following areas:

- **Data privacy and security**: This is the most common parental concern. Parents worry about their children's conversations being recorded, voices being saved, and personal information being accessed by third parties. Some parents also worry that data might be used for commercial purposes (e.g., targeted advertising).
- **Screen time**: Many parents worry that AI tools will increase children's "time staring at screens" and reduce interaction with real people. This concern is particularly prominent at the elementary and junior high levels.
- **Technology dependence**: Some parents worry that children will become overly dependent on AI, leading to a decline in independent thinking skills. For example, if children always use AI for translation, will they still work hard to learn vocabulary?
- **Content appropriateness**: Parents worry that AI might generate inappropriate content (e.g., violence, pornography, or age-inappropriate information). Although most AI tools have content filtering mechanisms, they are not perfect.
- **Socio-emotional development**: Some parents worry that if children become accustomed to conversing with AI, they may find it harder to form emotional connections with real people. Does the AI's "infinite patience" and "non-judgmental" quality affect children's expectations of real human relationships?

This study collected expressions of these concerns through parent interviews (week 2 and week 18) and analyzed how their concerns changed (or did not change)

over the semester. These data provide qualitative evidence for understanding the social acceptability of educational AI.

4.4 Implications of the Theoretical Review for This Study

The RRI framework and privacy principles provide theoretical tools for analyzing "teachers' ethical practices" and "parent narratives of concern" in this study:

The five dimensions of the RRI framework (anticipation, reflexivity, inclusion, responsiveness, sustainability) can be used to describe and analyze teachers' ethical decision-making in AI use.

The OECD privacy principles (minimal collection, transparency, data deletion, purpose limitation, parental consent, data security) provide concrete guidance for the design of the parent letter and data management in this study.

The classification of parental concerns (Akgun & Greenhow, 2022) provides a preliminary coding framework for analyzing parent interview data.

However, existing research is primarily based on theoretical reasoning or large-scale survey research, lacking qualitative descriptions of "how teachers actually handle ethical dilemmas in daily practice." This study attempts to fill this gap through the researcher's reflective journals and analysis of parent interviews.

5 Gaps in Existing Research and the Positioning of This Study

5.1 Major Gaps in Existing Research

Based on the literature review above, this study identifies the following major research gaps:

First, insufficient empirical research on human-machine negotiation of meaning. The Interaction Hypothesis and Output Hypothesis are primarily based on research on human-human interaction. Although a small number of studies have begun exploring learner-chatbot interaction (e.g., Tsai, 2021), these studies are mostly conducted in laboratory settings with short durations (typically 15–30 minutes per session) and lack qualitative descriptions of long-term interaction in real classrooms.

Second, absence of longitudinal tracking studies of AI role perception. Luckin et al.'s (2016) AI role framework is primarily based on theoretical reasoning. Although some studies have begun exploring learners' perceptions of AI (e.g., Godwin-Jones, 2022), there is a lack of tracking studies on how AI role perception evolves over a semester (or longer). Is students' perception of AI when they first encounter it the same as their perception after 18 weeks of use? If different, how does the change occur? These questions remain unanswered.

Third, limited research on privacy and ethics in educational AI focusing on classroom practice. Existing research has focused primarily on macro-level ethical principles (e.g., RRI framework, OECD privacy principles) or survey-based studies of parental concerns. However, "how teachers actually handle ethical dilemmas in daily teaching"—for example, what to do when a student inadvertently types personal information, or how to respond when AI generates inappropriate content—these micro-level ethical practices have not been adequately described.

Fourth, little research focusing on Chinese as a second language, junior high level, and international school contexts. Existing AI-assisted language learning research focuses primarily on English as a second language. Research on Chinese as a second language is very limited, and research focusing on the junior high level (ages 12–14) in international school contexts is almost non-existent. The psychological characteristics of junior high students (high anxiety, susceptibility to boredom, sensitivity to peer evaluation) differ from those of adults or elementary students, and the international school context (multicultural backgrounds, strong heterogeneity in proficiency levels) also differs from ordinary public schools. Therefore, research focusing on this specific context has unique value.

Fifth, despite the focus on Chinese as a second language (CSL), the existing literature on CSL in international school settings is surprisingly sparse. Researchers such as Duff (2008) have examined multilingual classroom dynamics in international schools, while Lam (2006) explored identity construction among adolescent Chinese learners. Pan and Block (2011) provided a critical sociolinguistic analysis of Chinese language teaching in multilingual contexts. More recently, Wu and colleagues have investigated technology-mediated Chinese instruction, including the use of chatbots and mobile apps for character learning and oral practice. The present study contributes to this underdeveloped area by providing thick description of CSL learners' interactions with AI in a specific international school context.

5.2 Theoretical Positioning of This Study

Based on the identification of gaps in existing research above, this study attempts to make contributions in the following four areas:

First, in terms of theoretical dialogue, this study does not simply "apply" SLA theories to the AI context but examines their applicability in human-machine interaction through qualitative observations in real classrooms and explores possible extensions. For example, does the AI's "predictable error" trigger output modification more effectively than the "natural errors" in human-human interaction? This question will be answered through analysis of classroom observation data.

Second, in terms of empirical contribution, this study provides one semester of longitudinal tracking data, describing the trajectory of changes in students' perceptions of AI roles. This is not only an empirical test of Luckin et al.'s (2016) framework but also a qualitative description of the concept of "role fluidity."

Third, in terms of ethics research, this study presents ethical challenges and coping strategies in everyday classroom practice of educational AI through analysis of teacher reflective journals and parent interviews, providing qualitative cases for the application of the RRI framework in the field of educational AI.

Fourth, in terms of context specificity, this study focuses on the specific context of a junior high CSL classroom in an international school, providing transferable experiences for teachers and researchers in similar contexts.

5.3 Relationship Between This Study and Existing Theories

The relationship between this study and existing theories can be described as "dialogue" rather than "application":

This study uses the Interaction Hypothesis, Output Hypothesis, and Affective Filter Hypothesis as conceptual tools for analyzing student behaviors, but does not presume that these hypotheses "necessarily hold" in human-machine interaction. Instead, this study remains open to the possibility of "discovering different patterns."

This study borrows Luckin et al.'s (2016) AI role framework as a starting point for preliminary classification, but allows new role types or role dynamics (e.g., "role fluidity") to emerge from the data.

This study adopts the RRI framework and OECD privacy principles as reference frameworks for analyzing teachers' ethical practices, but focuses on their practical operation and challenges in real classrooms.

This "dialogue" rather than "application" positioning is characteristic of the qualitative research paradigm—theory is not used for "verification" but for dialogue, mutual questioning, and mutual inspiration with data.

6 Chapter Summary

This chapter has systematically reviewed theoretical literature relevant to the three core research questions of this study:

In the field of second language acquisition, the Interaction Hypothesis (Long, 1996), Output Hypothesis (Swain, 2005), and Affective Filter Hypothesis (Krashen, 1982) provide conceptual tools for understanding "human-machine negotiation of meaning" in this study. The common concern of these theories is how learners acquire language through negotiation of meaning and output modification in interaction, and how affective states affect this process.

In the field of computer-assisted language learning, Luckin et al.'s (2016) AI role framework (tutor, coach, peer, tool) provides preliminary classification for analyzing "students' perceptions of AI roles" in this study. Extensions by subsequent researchers—e.g., Godwin-Jones's (2022) concept of role fluidity and Tsai's (2021) "AI as a mistake-prone peer"—provide an empirical foundation for dialogue.

In the field of educational AI privacy and ethics, the Responsible Research and Innovation framework (von Schomberg, 2013) and the OECD (2022) privacy principles provide theoretical tools for analyzing "teachers' ethical practices" and "parent narratives of concern" in this study. Akgun and Greenhow's (2022) review of parental concerns provides coding references for the parent interviews in this study.

Major gaps in existing research include insufficient empirical research on human-machine negotiation of meaning, absence of longitudinal tracking studies of AI role perception, limited research on educational AI ethics focusing on classroom practice, and little research focusing on CSL junior high international school contexts. Based on these gaps, this study has established its theoretical positioning as "dialogue rather than application."

Chapter 3
Research Methodology

Huang Xiao

Abstract: This chapter details the research design, participants, data collection methods, data analysis methods, the researcher's role and reflexivity, and ethical considerations of this study. This study adopts a qualitative research paradigm, using a single-case longitudinal design to track the entire process of AI-assisted learning over 18 weeks in a junior high Chinese as a second language (CSL) classroom at an international school.

1 Introduction

The structure of this chapter is as follows: first, the rationale for choosing the qualitative research paradigm and the characteristics of the single-case longitudinal design are explained; second, the research field and participant details are described; third, six data collection methods (classroom observations and video recordings, dialogue logs, student task cards, semi-structured interviews, open-ended parent communication, and teacher reflective journals) are introduced individually; fourth, the specific steps of thematic analysis and the analytical framework are elaborated; finally, the researcher's dual role, the trustworthiness of the study, and ethical considerations are discussed.

Huang Xiao • Krirk University, Bangkok, Thailand
e-mail: huang.xiao@krirk.ac.th

© The Author(s). Published in the USA.
https://doi.org/10.63408/979-8-9928364-4-8

2 Research Paradigm and Design

2.1 Rationale for Choosing the Qualitative Research Paradigm

This study adopts a qualitative research paradigm. The rationale for choosing the qualitative paradigm is based on the following three considerations:

First, the nature of the research questions. The three research questions proposed in this study—"What meaning negotiation behaviors occur?", "How do students perceive the AI's role?", and "How does the teacher handle ethical dilemmas?"—are all questions about "meaning," "perception," and "process," rather than about "how much," "how often," or "correlation." Qualitative research excels at answering "what," "how," and "why" questions and can delve into specific contexts to understand participants' perspectives and the complexity of their behaviors (Denzin & Lincoln, 2018).

Second, the specificity of the research context. AI-assisted language teaching is an emerging and rapidly changing field. Human-machine interaction in the classroom is influenced by multiple factors—students' language proficiency, personality, mood of the day, AI's technical performance, physical classroom environment, and so on. These factors are intertwined and difficult to "control" or "isolate" through quantitative design. Qualitative research allows researchers to observe these complex phenomena in natural settings rather than simplifying them in a laboratory.

Third, the purpose of the research. The purpose of this study is not to "prove" that AI is effective or ineffective, nor to "measure" the effect of a variable. This study aims to provide rich descriptions and in-depth understanding of the process of AI use in real classrooms, laying a foundation for subsequent research and providing references for teaching practice. This exploratory and descriptive purpose is highly consistent with the orientation of the qualitative research paradigm.

A brief elaboration on the term "Generative CALL" introduced in Chapter 2 is warranted here. The author suggests that the emergence of generative AI (e.g., ChatGPT) may constitute a fourth stage in CALL's evolution, following Warschauer and Healey's (1998) three-stage model (Behavioristic → Communicative → Integrative). What distinguishes Generative CALL from previous stages is not merely technological sophistication but a qualitative shift in the human-computer relationship. In earlier stages, the computer executed pre-programmed responses; in Generative CALL, the AI generates novel, context-sensitive responses and can even take on personas (e.g., "a directionally challenged fool"). This shift has pedagogical implications: the AI is no longer a tool that merely responds but a partner that can be "taught," "corrected," and "negotiated with." However, this study does not claim that Generative CALL represents a universal or inevitable stage; rather, it is offered as a heuristic for understanding the current moment. Future historical analyses may revise or reject this periodization.

2.2 Single-Case Longitudinal Design

This study adopts a single-case longitudinal design. Yin (2018) points out that case study research is suitable for situations where the researcher asks "how" and "why" questions about a contemporary phenomenon over which the researcher has little control. The "case" in this study is the entire process of AI-assisted learning over 18 weeks in a junior high CSL classroom (14 students) at an international school.

The rationale for choosing a single case rather than multiple cases is that this study pursues depth rather than breadth. Through long-term observation in one classroom, the researcher can build trusting relationships with students, track the trajectories of individual students' changes, and capture subtle but important interactional details in the classroom. Multiple-case studies might enhance the "generalizability" of findings but would sacrifice the depth of description and richness of context.

"Longitudinal" means that data collection covers the entire semester (18 weeks). This design allows this study to:

- Track the evolution of students' perceptions of AI roles (rather than recording a snapshot at a single point in time);
- Observe changes in students' meaning negotiation behaviors (from first contact with AI to proficient use);
- Collect pre-post changes in parent attitudes (beginning and end of semester).

2.3 Statement of the Researcher's Role

In this study, the researcher simultaneously serves as the classroom's Chinese teacher and curriculum designer. This "teacher-researcher" dual role is both an advantage of this study and a challenge that requires careful handling.

Advantages:

- As the teacher, the researcher has deep access to the field and can build trust with students within the natural teacher-student relationship;
- As the curriculum designer, the researcher has a clear understanding of how AI tools are used and the instructional objectives of each activity, allowing comparison between classroom observations and design intentions;
- Long-term presence (the entire semester) enables the researcher to capture everyday interactional details that "one-time visitors" cannot observe.

Challenges:

- Students may give the answers "the teacher wants to hear" in interviews because the researcher is also their teacher;
- The researcher may have "emotional investment" in their own curriculum design, affecting the objectivity of observation and interpretation;

- The researcher's instructional decisions (e.g., adjusting activity time, changing AI roles) are themselves part of the research, which may make it difficult to distinguish the research process from daily teaching.

Strategies for addressing these challenges are detailed in Sections 3.8 (Researcher's Reflexivity) and 3.9 (Trustworthiness).

3 Research Field and Participants

3.1 School and Classroom Context

This study was conducted at an international school. The school is located in Bangkok and is a K-12 international school offering the International Baccalaureate (IB) curriculum. The student body has diverse cultural and linguistic backgrounds. The language of instruction is English, but Chinese as a second language is a compulsory subject.

The study was conducted in a CSL classroom in the junior high division. This class has four 50-minute Chinese lessons per week. One lesson per week (typically on Friday) was designated as an "AI-assisted lesson," following the 18-week curriculum outline described in Section 1.1.3. The remaining three lessons were traditional Chinese lessons (without AI intervention), including vocabulary writing, grammar instruction, pair dialogue practice, cultural activities, etc.

The classroom is approximately 50 square meters, equipped with an electronic whiteboard, projector, and WiFi coverage. Students use school-provided tablets (one per student) during class, pre-installed with applications including ChatGPT, Duolingo, and Canva. The school provided institutional access to ChatGPT Plus (GPT-4.0, voice mode) for all students at no personal cost. A tripod camera was set up in the back corner of the classroom for video recording (with student informed consent).

3.2 Student Participants

The class had 14 students. Their basic information is as follows (Table 3):

Table 3 Student participant characteristics

Characteristic	Description
Age range	12-14 years old
Grade	Junior high Grade 2 (equivalent to Grade 8)
Chinese proficiency	HSK Level 3 (approximately 600 vocabulary words, capable of daily communication)
Length of Chinese study	1.5-3 years (mean 2.1 years)
Gender	8 female, 6 male
Native language background	English (6), Korean (3), Japanese (2), French (1), German (1), Spanish (1)

All students received research information and informed consent forms at the beginning of the semester. Parents signed consent forms, and students themselves verbally agreed to participate (a collective explanation and individual confirmation were conducted in class during Week 1). Students were informed that participation was entirely voluntary, that they could withdraw at any time, and that withdrawal would not affect their Chinese course grades.

To protect participant privacy, this study uses pseudonyms (S1-S14) to refer to students. The selection of focal students (S3, S7, S11) was based on purposive sampling (Patton, 2015)—selecting students with different Chinese proficiency levels (high, intermediate, basic) and different genders and native language backgrounds to capture diverse perspectives. S3 (female, native English, intermediate-high Chinese) represents a "stable learner"; S7 (male, native Korean, weak Chinese foundation) represents a "learner transitioning from anxiety to active participation"; S11 (female, native French, high Chinese proficiency) represents an "efficient but emotionally less invested learner."

3.3 Parent Participants

Parents of all 14 students received a parent information letter (see Appendix A) at the beginning of the semester. Among them, parents of the six focal students (parents corresponding to S3, S7, S11) agreed to participate in interviews. Additionally, at the end of the semester, all parents received an open-ended questionnaire (see Appendix B), and nine parents returned the questionnaire.

4 Data Collection Methods

This study employs a multi-source qualitative data collection strategy, using six data sources to achieve triangulation. The purpose, collection frequency, and specific content of each data source are described below.

4.1 Classroom Observations and Video Recordings

Purpose: To capture real-time behaviors of students interacting with the AI, including language output, non-verbal behaviors (gestures, facial expressions), and interactions with peers.

Collection frequency: Once per week (AI-assisted lessons), for a total of 18 times. Among these, weeks 1, 5, 9, 13, and 17 included full video recordings (whole-class panoramic view + focal student screen recordings); the remaining weeks included only field observation notes (without video, to reduce impact on students).

Procedure:

- Before class, set up a tripod camera at the back of the classroom to record the whole-class view. Activate screen recording on the tablets of focal students (S3, S7, S11) to capture complete dialogues between students and the AI (including student voice, AI voice, and on-screen actions).
- During class, the researcher (as teacher) takes real-time handwritten field notes while circulating, focusing on:
  - Volume, tone, and pauses when students speak to the AI;
  - Students' gestures, facial expressions, and body orientation;
  - Whether students turn to peers for help or comments;
  - Whether students exhibit emotions such as anxiety, pleasure, or frustration.
- After class, the researcher watches the recordings, expands field notes into detailed narrative records, and transcribes key dialogue segments between focal students and the AI verbatim. Observation record forms: Use the "Meaning Negotiation Event Recording Form" and "AI Role Perception Recording Form" designed in Section 3.4.1 (see Appendix C), completed in narrative format.

4.2 Student-AI Dialogue Logs

Purpose: To obtain precise written records of student-AI dialogues for analyzing linguistic details of meaning negotiation.

Collection frequency: Exported after each AI activity, for a total of 18 times.

Procedure:

- After each AI activity, the researcher collects screenshots of ChatGPT dialogue histories from the tablets of focal students (S3, S7, S11).
- Dialogue histories include: each utterance input by the student, each response from the AI, and timestamps.
- The researcher transcribes the dialogue histories verbatim into analysis documents, preserving the original format (including students' spelling errors, punctuation, and mixing of Chinese and English).

Privacy protection: All dialogue logs are deleted from tablets immediately after transcription. Pseudonyms are used in transcription documents, and no identifiable personal information is included.

4.3 Student Task Cards

Purpose: To collect students' handwritten drafts "before AI interaction" and revision traces "after interaction" to understand students' output modification processes.

Collection frequency: Once per week (task cards from AI-assisted lessons), for a total of 18 times.

Task card content (using Week 5 as an example, see Appendix D):

- Before conversing with the AI, students handwrite planned output sentences (e.g., route descriptions) on the task card.
- During the AI dialogue, students record the AI's responses and their own revisions on the task card.
- After the activity, students check self-assessment items on the task card ("Did the AI understand?", "How many revisions did I make?").

Analytical focus:

- Differences between handwritten drafts and final output;
- Number of revisions and revision methods recorded by students themselves;
- "Reasons for revision" written by students (e.g., "The AI didn't understand 'traffic light,' so I added the English word").

4.4 Semi-Structured Interviews

Purpose: To gain in-depth understanding of students' perceptions of the AI, affective experiences, and awareness of privacy issues.

Interviewees: Six focal students (S3, S7, S11, S5, S9, S13—with S5, S9, S13 as "alternate focal students" to be used if S3/S7/S11 are absent or decline interviews), all student identifiers follow the S1-S14 format throughout the manuscript.

Interview timing: Week 6 (mid-semester) and Week 18 (end of semester), each lasting 20-25 minutes.

Interview format: One-on-one, conducted in a quiet corner outside the classroom, audio-recorded in full.

Interview protocol: See Appendix E. The protocol includes four parts:

- Experience with AI dialogue ("Has the AI ever failed to understand you? What did you do?")
- Perception of AI's role ("What kind of presence do you think the AI is like?")
- Anxiety and affective experience ("Do you feel nervous speaking to the AI? What about speaking to classmates?")
- Privacy and security perception ("Do you know that the AI records what you say? Does that worry you?")

Audio processing: Interview recordings were transcribed verbatim by the researcher. Transcripts were returned to students for confirmation (member checking) before analysis.

4.5 Open-Ended Parent Communication

Purpose: To understand parents' attitudes toward the introduction of AI into the Chinese classroom, their concerns, and changes they observe in their children.

Collection methods: Including two interviews (with parents of focal students, Week 2 and Week 18) and two open-ended questionnaires (all parents, Week 1 and Week 18).

Parent interviews (parents of focal students P3, P7, P11, 15-20 minutes each):

- Week 2 interview focus: Initial attitudes, main concerns, feedback on the parent letter.
- Week 18 interview focus: Changes in attitudes, observed changes in children, suggestions for next semester.

Parent open-ended questionnaires (parents of all 14 students, Week 1 and Week 18):

- Week 1 questionnaire (distributed with the parent letter): "What are your main concerns about introducing AI into the Chinese classroom?" "What would you like the teacher to pay attention to?"
- Week 18 questionnaire (distributed at the end of the semester): "Have your attitudes toward AI-assisted Chinese lessons changed over this semester?" "Have you observed your child mentioning the AI or Chinese lessons at home?"

Questionnaire return rate: 12 returned in Week 1, 9 returned in Week 18.

4.6 Teacher Reflective Journals

Purpose: To document the researcher's (as teacher) decisions, confusions, unexpected events, and handling of ethical dilemmas during AI use.

Collection frequency: Written after each week's lesson, for a total of 18 times.

Journal content (narrative format, not a checklist):

- Operation of the AI tool in that lesson (any technical issues? student reactions?);
- Reflections on instructional decisions ("Today I shortened the AI dialogue time from 15 minutes to 10 minutes because..."");
- Documentation of ethical dilemmas ("A student inadvertently typed their full name in the dialogue; I immediately cleared the record and reminded the whole class..."");
- Supplementary observations of focal students ("S7 was particularly active today, possibly because..."").

Journal length: Approximately 500-1000 words per entry.

5 Data Collection Timeline

Week	Data Collection Activities
Week 1	Classroom observation (video + screen recording); student task cards; teacher reflective journal; parent questionnaire (Round 1) distributed
Week 2	Classroom observation (field notes); parent interviews (P3, P7, P11)
Weeks 3-4	Classroom observation (field notes); student task cards; teacher reflective journal
Week 5	Classroom observation (video + screen recording); student task cards; teacher reflective journal
Week 6	Student interviews (S3, S7, S11, mid-semester); teacher reflective journal
Weeks 7-8	Classroom observation (field notes); student task cards; teacher reflective journal
Week 9	Classroom observation (video + screen recording); student task cards; teacher reflective journal
Weeks 10-12	Classroom observation (field notes); student task cards; teacher reflective journal
Week 13	Classroom observation (video + screen recording); student task cards; teacher reflective journal
Weeks 14-16	Classroom observation (field notes); student task cards; teacher reflective journal
Week 17	Classroom observation (video + screen recording); student task cards; teacher reflective journal
Week 18	Student interviews (S3, S7, S11, end of semester); parent interviews (P3, P7, P11); parent questionnaire (Round 2); teacher reflective journal

6 Data Analysis Methods

6.1 Thematic Analysis

This study employs Thematic Analysis (Braun & Clarke, 2006, 2022) as the data analysis method. Thematic analysis is a qualitative analytic method for identifying, analyzing, and reporting patterns (themes) within data. Its advantage lies in its flexibility—it is not bound to specific theoretical frameworks and can be used across various qualitative research paradigms. Braun and Clarke (2006) proposed a six-phase process for thematic analysis. The specific operations for these six phases in this study are as follows:

Phase 1: Familiarizing oneself with the data The researcher reads all data repeatedly—classroom observation records, dialogue logs, task cards, interview transcripts, teacher reflective journals—until achieving an in-depth understanding of the overall content. During this phase, the researcher simultaneously makes preliminary notes (not coding, but recording impressions and questions). Specific operations:

- Import all data into NVivo qualitative analysis software;
- Read data by type (e.g., all observation records first, then all interview transcripts);
- After reading each piece of data, write a brief "initial impressions" note.

Phase 2: Generating initial codes The researcher codes the data line by line, marking segments relevant to the research questions. Codes are "the smallest units of meaning in the data"—they can be a word, a sentence, or a paragraph. This study uses a hybrid coding strategy (Fereday & Muir-Cochrane, 2006):

- Theory-driven coding: Preset initial codes based on concepts from the literature review (e.g., "negotiation of meaning," "output modification," "affective filter," "AI role");
- Data-driven coding: Allow new codes to emerge from the data while reading (e.g., "it's so silly," "the process of teaching AI is also learning").

Table 4 Example codes (from classroom observation records)

Data Segment	Code
"No! Not turn left. Go straight first!"	negotiation_correction
"It's so silly" (laughing, to peer)	AI_as_fallible_peer + affective_positive
"Traffic light is 'honglüdeng'"	negotiation_paraphrase + code_switching
"When I speak to the AI, it won't laugh at me even if I make mistakes"	affective_low_anxiety

Phase 3: Searching for themes The researcher groups similar codes together to form initial "candidate themes." The goal of this phase is to move from a "list of codes" to an "organized thematic structure." Operations:

- In NVivo, group all codes by similarity;
- Write a "theme description" for each code group (what does this theme say?);
- Examine logical relationships among themes (coordinate, hierarchical, causal, etc.).

Table 5 Example (from codes to themes)

Codes	Theme
negotiation_correction, negotiation_paraphrase, negotiation_repetition	"Ways students correct the AI"
AI_as_fallible_peer, AI_as_opponent, affective_positive	"AI as a mistake-prone peer: positive affective experience"
AI_as_tool, affective_neutral	"AI as a tool: neutral affective experience"

Phase 4: Reviewing themes The researcher reviews candidate themes, checking:

- Whether themes are internally coherent (do the codes truly belong to the same theme?);
- Whether themes are clearly distinguishable (are different themes truly different?);
- Whether themes have sufficient data support (not just one or two codes?).

After review, the researcher may merge similar themes, split overly broad themes, or delete themes with insufficient data support.

Phase 5: Defining and naming themes The researcher writes precise definitions for each theme, including:

- Core content of the theme ("What phenomenon does this theme describe?");
- Boundaries of the theme ("What does not belong to this theme?");
- Typical evidence for the theme (illustrated with a representative data segment).

Table 6 Example theme

Theme Name	"From the corrected to the corrector: role reversal brought by AI's mistakes"
Definition	This theme describes the process in which, when the AI deliberately makes mistakes, students shift from "fearing being judged" to "actively judging the AI." In this process, students experience a sense of competence, anxiety levels decrease, and engagement increases.
Boundaries	This theme does not include situations where students "calmly correct the AI" (that belongs to another theme, "AI as a tool"), but only includes correction behaviors accompanied by positive affective experiences (laughter, excitement, comments to peers).
Typical Evidence	S7, in Week 5, laughed and said to a peer, "It's so silly," then loudly corrected the AI's directional error.

Phase 6: Producing the report The researcher organizes themes into coherent narratives, presented in Chapters 4, 5, and 6 of the thesis. In writing, each theme is accompanied by: theme description; representative data segments (verbatim student quotes, narrative observation records); connections to research questions; dialogue with theory.

6.2 Analytical Framework: Theory-Data Dialogue

The analysis in this study is not "applying theory" to data but rather a theory-data dialogue. Specifically, the researcher adheres to the following principles:

Principle 1: Theory as "lens" rather than "mold." The theories from the literature review (Interaction Hypothesis, Output Hypothesis, AI role framework, RRI framework) provide initial "lenses" for analysis—they help the researcher notice certain phenomena (e.g., meaning negotiation, output modification, role perception). However, the researcher does not presuppose that "data must fit theory." If data are inconsistent with theory (e.g., student-AI interaction patterns differ from human-human interaction described by Long), the researcher reports this honestly and explores possible reasons.

Principle 2: Data take precedence over theory. In the process of coding and searching for themes, the researcher first "listens to the data," allowing themes to emerge from the data. Theory-driven coding serves only as a supplement (e.g., using "negotiation of meaning" as an initial code) but cannot "override" other important patterns appearing in the data (e.g., the affective expression "it's so silly").

Principle 3: Active search for disconfirming evidence. The researcher actively seeks data contrary to dominant themes (negative cases). For example, when finding that "most students exhibit positive affect when the AI makes mistakes," the researcher asks: Are there students who exhibit frustration when the AI makes mistakes? If so, why? Analysis of these negative cases is included in the research findings to enhance depth and credibility.

6.3 Analysis Software

This study uses NVivo 14 qualitative analysis software to assist with data management, coding, and theme organization. NVivo functions include:

- Importing multiple data formats (text, PDF, audio, video);
- Coding directly on data;
- Organizing codes into hierarchical structures (tree nodes);
- Searching and retrieving coded segments;
- Visualizing relationships among themes.

The researcher uses NVivo's "memo" function to record reflections and decisions during the analysis process, creating an "analysis trail" to enhance research transparency.

A note on the use of qualitative analysis software: Some scholars express concern that CAQDAS (Computer-Assisted Qualitative Data Analysis Software) like NVivo can "mechanize" what should be an interpretive, iterative process—reducing qualitative analysis to code-and-retrieve operations. This study acknowledges that concern and took deliberate steps to maintain interpretive primacy. NVivo was used as a data management tool, not an analytical engine. The researcher completed initial line-by-line coding on printed transcripts before entering codes into NVivo. Thematic development (moving from codes to candidate themes) was done manually using concept maps and sticky notes, not through NVivo's automated theme-finding functions. NVivo's primary contributions were: (a) efficient retrieval of coded segments across multiple data sources, (b) tracking of code frequencies to identify saturation points, and (c) memo integration to maintain an audit trail. In all cases, the researcher's interpretive judgment—not software output—determined final themes.

7 Trustworthiness of the Study

In qualitative research, the concepts of "reliability" and "validity" from quantitative research are replaced by a trustworthiness framework (Lincoln & Guba, 1985). Lincoln and Guba proposed four qualitative research quality criteria: credibility, transferability, dependability, and confirmability. The strategies for achieving these four criteria in this study are as follows:

7.1 Credibility

Credibility corresponds to "internal validity" in quantitative research, referring to whether the research findings truly reflect participants' experiences and perspectives. Strategies in this study:

- Prolonged engagement: The researcher conducted 18 weeks of continuous observation in one classroom, building trusting relationships with students and reducing the impact of "researcher presence" on student behavior.
- Persistent observation: The researcher conducted weekly classroom observations, capturing patterns and changes that one-time observations cannot capture.
- Triangulation: Six data sources (observations, logs, task cards, interviews, parent questionnaires, reflective journals) were used to cross-validate findings.
- Member checking: Interview transcripts were returned to students for confirmation; preliminary analysis results (themes) were shared with focal students at the end of the semester, asking "Have I understood correctly? Is anything missing?" Conducting member checking with 12-14-year-olds presented specific

challenges. First, social desirability bias: students may say "yes" to please the teacher-researcher even if they disagree. To mitigate this, the researcher emphasized (both verbally and in writing) that disagreement would be helpful, not a problem. Second, limited metalinguistic vocabulary: younger adolescents may lack the words to articulate disagreements with interpretations. The researcher provided concrete examples (e.g., "Do you think the AI is more like a tool or more like a friend?") rather than asking abstractly "Is my interpretation correct?" Third, attention span: students reviewed transcripts and themes in short 10-minute sessions rather than all at once. Despite these challenges, three focal students (S3, S7, and S11) offered substantive feedback—for example, S11 clarified that she saw the AI as a "tool" not a "coach," leading the researcher to adjust the coding of her interviews.

- Peer debriefing: A colleague not involved in this study (another Chinese teacher at the same school) was invited to review codes and themes, asking questions and proposing alternative interpretations.

7.2 Transferability

Transferability corresponds to "external validity" in quantitative research, referring to whether research findings can be transferred to other contexts. Qualitative research does not pursue "statistical generalization" but rather "theoretical generalization" and "contextual resonance." Strategies in this study:

- Thick description: Provide sufficiently detailed descriptions of classroom context, student characteristics, and curriculum design in the thesis, enabling readers to judge "whether this study's context is similar to my context."
- Specification of contextual boundaries: Clearly specify the boundary conditions of this study (junior high, international school, HSK Level 3, CSL, ChatGPT voice), avoiding overgeneralization.

Beyond Lincoln and Guba's (1985) framework, Stake (1995) offers the concept of "naturalistic generalization"—the recognition that readers of case study research naturally generalize, not statistically but intuitively, by recognizing similarities between the reported case and their own contexts. Naturalistic generalization does not require the researcher to claim representativeness; it relies on readers' practical wisdom. The implication for this study is that readers—other CSL teachers in international schools, for example—will judge for themselves whether S7's trajectory from anxious silence to excited correction to calm strategic use resonates with students in their own classrooms. To facilitate such naturalistic generalization, this study has provided "thick description" (Geertz, 1973) of the classroom context, student characteristics, curriculum design, and interactional details. Readers are encouraged to ask: "How similar is this context to mine? How dissimilar? What would I keep the same, and what would I change?" This is the proper use of qualitative case study

research: not as a source of universal laws, but as a source of vicarious experience and practical wisdom.

7.3 Dependability

Dependability corresponds to "reliability" in quantitative research, referring to whether the research process is consistent and traceable. Strategies in this study:

- Research trail documentation: Keep records of all research decisions (e.g., why this focal student was chosen, why the activity time was adjusted in Week 5), forming a "research decision log."
- Coding consistency check: Recode the same data after a one-month interval and compare consistency between the two coding sessions; invite a peer researcher to independently code a segment and compare coding differences.
- Raw data preservation: All data (recordings, transcripts, task card scans, interview audio) are properly preserved and available for review.

7.4 Confirmability

Confirmability corresponds to "objectivity" in quantitative research, referring to whether research findings derive from data rather than researcher bias. Strategies in this study:

- Reflexivity journal: The researcher writes a weekly reflexivity journal, recording their own assumptions, emotions, interactions with students, and how these factors may affect data collection and analysis (see Section 3.8).
- Audit trail: Preserve all analytical records from raw data to codes to themes to findings, enabling readers to trace "how the researcher arrived at this conclusion from the data."
- Negative case analysis: Actively report data contrary to dominant themes, avoiding "reporting only data that support one's own views."

8 Researcher's Reflexivity

Reflexivity is the continuous reflection by the qualitative researcher on how their own role, assumptions, and biases may affect the research process (Berger, 2015). This study addresses reflexivity in the following three areas:

8.1 Managing the Teacher-Researcher Dual Role

As noted earlier, the researcher simultaneously serves as teacher and researcher. This dual role creates the following tensions:

Tension	Researcher's Coping Strategy
Students may give answers "the teacher wants to hear"	Clearly state in interviews that "this interview is not a test; you can say anything you really think, including not liking the AI"; use neutral probes (e.g., "Can you give an example?") rather than leading questions (e.g., "You found the AI helpful, right?")
Researcher has emotional investment in own curriculum design	Actively seek disconfirming evidence in analysis (e.g., students disliking the AI, technical failures causing activity breakdowns); record in reflexivity journal: "Today I saw S7 very happy, and I felt good-could this emotion lead me to overemphasize positive cases?"
Instructional decisions and research decisions are difficult to separate	Clearly distinguish between "teaching log" (recording instructional decisions) and "research log" (recording research decisions); transparently report the researcher's dual role and its potential impact in the thesis

8.2 Researcher's Pre-understandings and Assumptions

Before entering the study, the researcher held the following pre-understandings:

- AI has the potential to reduce anxiety in language learning (based on personal teaching experience and literature reading);
- "Having AI make mistakes" may have greater pedagogical value than "pursuing a perfect AI" (based on Tsai's, 2021 research);
- Parents may be skeptical about introducing AI into the classroom (based on informal conversations with colleagues).

Coping strategy: The researcher explicitly records these pre-understandings in the reflexivity journal and actively asks during analysis: "Did I discover this code/theme from the data, or did I project my own assumptions onto the data?" Peer debriefing is also an important mechanism for checking bias.

8.3 Managing Emotional Responses

During long-term interaction with students, the researcher developed positive feelings toward certain students (e.g., S7, who transitioned from anxiety to active par-

ticipation). These feelings may affect the researcher's interpretation of data—for example, unconsciously "glorifying" S7's performance.

Coping strategies:

- Record emotional responses to each focal student in the reflexivity journal;
- During analysis, attend to both S7's "high moments" and "frustration moments";
- Invite a peer researcher to independently analyze S7's data and compare whether interpretations are consistent.

9 Ethical Considerations

9.1 Informed Consent

Students: In the first week of the semester, the researcher used 10 minutes to explain the purpose, procedures, and rights of the study to the whole class, using language understandable to students (English + simple Chinese). Students were informed that:

- Participation was entirely voluntary, and they could withdraw at any time;
- Withdrawal would not affect their Chinese course grades;
- Classroom recordings and audio were for research purposes only and would not be made public;
- Pseudonyms would be used in the thesis, and no one would be able to identify them.

After all students verbally agreed, the researcher distributed a "Student Informed Consent Form" (to be signed by parents, as students were minors).

Parents: At the beginning of the semester, the researcher sent a "Letter to Parents" (see Appendix A) to all parents, explaining in detail:

- Why AI is used in Chinese lessons;
- Which AI tools are used;
- Privacy protection measures (no recording, no saving, use of pseudonyms, etc.);
- Parents' rights (opt-out, ask questions at any time).

After parents signed the consent form, their children could participate in AI activities. In this study, two parents initially expressed concerns; the researcher conducted phone conversations with them, explained the specific measures, and they agreed to their children's participation.

9.2 Privacy and Data Protection

This study adopted the following privacy protection measures:

Data Type	Protection Measures
Classroom video recordings	Viewed only by the researcher; stored on an encrypted hard drive after analysis; destroyed 12 months after study completion
Dialogue logs	Deleted from tablets immediately after export; pseudonyms used in transcription files
Student task cards	Scanned and stored in encrypted folders; paper versions destroyed after scanning
Interview audio recordings	Deleted immediately after transcription; any identifiable information removed from transcripts
Parent questionnaires	Completed anonymously (no name required); questionnaire data stored separately from research data

Handling of special situations: In Week 4, during a classroom activity, one student (S8) inadvertently typed their full name in the dialogue. The researcher immediately:

- Cleared that dialogue history (deleted it in ChatGPT);
- Reminded the whole class after class: "Do not type your name in the AI";
- Documented this event in the teacher reflective journal and reflected on "Why did the student do this? Do I need to provide clearer reminders?"

9.3 Minimizing Harm and Benefits

Potential risks to students in this study include: frustration due to technical failures, anxiety when AI interaction does not go smoothly, and (a very small possibility) the AI generating inappropriate content.

The researcher's coping strategies:

- Each AI lesson had a "no-AI backup activity" ready for immediate switch in case of technical failure (e.g., pair-based role-play using printed dialogue cards);
- During teacher circulation, close attention was paid to students' affective states; if any student showed persistent frustration, immediate intervention was provided;
- All AI tools used were educational versions or had content filtering mechanisms; the researcher pre-tested all AI activities.

Potential benefits of this study include: students gaining more oral practice opportunities, experiencing success in "speaking Chinese" in a low-anxiety environment, and developing AI literacy (an important 21st-century skill).

10 Chapter Summary

This chapter has detailed the research methodology of this study. This study adopts a qualitative research paradigm, using a single-case longitudinal design to track the entire process of AI-assisted learning over 18 weeks in a junior high CSL classroom at an international school. Data collection used six sources: classroom observations and video recordings, student-AI dialogue logs, student task cards, semi-structured interviews, open-ended parent communication, and teacher reflective journals. These multiple data sources enabled triangulation, enhancing the trustworthiness and richness of the research findings.

Data analysis employed Braun and Clarke's (2006, 2022) thematic analysis, following the six-phase process: familiarizing with data, generating initial codes, searching for themes, reviewing themes, defining and naming themes, and producing the report. The analysis adopted a "theory-data dialogue" principle, using theory as a lens while allowing data to challenge theory. Trustworthiness was ensured through prolonged engagement, triangulation, member checking, peer debriefing, thick description, research trail documentation, reflexivity journals, and audit trails. Ethical considerations covered informed consent, privacy protection, harm minimization, and institutional approval. The researcher paid particular attention to the tensions arising from the "teacher-researcher" dual role and addressed this challenge through reflexivity journals and peer debriefing.

Chapter 4
Research Findings I – Human-Machine Negotiation of Meaning

Huang Xiao

Abstract: This chapter presents the first core finding of this study, addressing Research Question RQ1: In an AI-assisted junior high CSL classroom, what kinds of meaning negotiation behaviors occur between students and the AI? How does the AI's "intentional error-making" setting affect students' language output and modification?

1 Introduction

Based on thematic analysis of 18 weeks of classroom observation records, student-AI dialogue logs, student task cards, and focal student interview data, this chapter extracts three main themes:

- **Theme 1: "It Doesn't Understand" - Triggers of Meaning Negotiation and Students' Initial Responses.** This theme describes students' first reaction patterns when the AI "fails to understand" or "makes a mistake," and how these patterns change over time.
- **Theme 2: "Let Me Teach You" - Students' Meaning Negotiation Strategies as "Correctors."** This theme describes the specific strategies students use to overcome comprehension obstacles with the AI, including repetition, paraphrase, simplification, code-switching to English, and multimodal assistance, and analyzes the effectiveness of these strategies.

Huang Xiao • Krirk University, Bangkok, Thailand
e-mail: huang.xiao@krirk.ac.th

© The Author(s). Published in the USA.
https://doi.org/10.63408/979-8-9928364-4-8

- **Theme 3: "The Process of Teaching It Is Also Learning" - The Impact of Meaning Negotiation on Language Output.** This theme explores how meaning negotiation triggers students' noticing, hypothesis testing, and output modification, and how students perceive the value of this process for their Chinese learning.

Each theme is supported primarily by narrative segments from classroom observations and verbatim student quotes, supplemented by observation notes from the researcher's reflective journals. The presentation in this chapter is descriptive and contextualized, striving to enable readers to "see" the authentic interactions occurring in the classroom. The roadmap could be described as: Theme 1 (Triggers of meaning negotiation) → Theme 2 (Students' negotiation strategies) → Theme 3 (Impact on language output).

2 Theme 1: "It Doesn't Understand" - Triggers of Meaning Negotiation and Responses

2.1 The Al's "Failure to Understand" as a Trigger for Negotiation

In the curriculum design of this study, the AI was intentionally set to be a "mistake-prone" character. Whether it was the "silly little robot" in Weeks 1-6 or the specially designed "directionally challenged Xiao A" in Week 5, the AI actively created comprehension obstacles during dialogue-it would say "I don't quite understand" to students' perfectly correct sentences, deliberately give wrong directions, and confuse key information during confirmation checks. The pedagogical intention behind this "intentional error-making" design was to create opportunities for meaning negotiation. In the traditional "perfect AI" model, the AI always gives correct responses, and students merely "input-receive correct output," requiring almost no meaning negotiation. The design of this study was precisely the opposite the Al's "silliness" forced students to pause during dialogue to explain, repeat, modify, and correct. Based on 18 weeks of observations, the Al's "failure to understand" did effectively trigger meaning negotiation. In every Al lesson, nearly all focal students experienced at least one comprehension obstacle with the AI and engaged in corresponding interactional adjustments. Among the meaning negotiation events observed by the researcher, the vast majority were initiated by the AI (through utterances such as "I don't understand" or "Do you mean...?"), with a few initiated by students (when students noticed the AI had given a clearly erroneous response). However, students' initial responses to the Al's "failure to understand" were not static. From Week 1 to Week 18, the researcher observed a clear trajectory of change.

2.2 Initial Stage (Weeks 1-4): Tension, Silence, and Tentativeness

In Weeks 1-4, when the AI first said “I don’t quite understand,” most students’ first reactions were tension and silence. The following is a typical classroom observation record (Week 1, S7, first conversation with the AI):

> S7 said to the tablet: “My dad is tall.”
> AI: “Huh? What does ‘tall’ mean? I don’t quite understand.”
> S7 froze. He did not respond to the AI immediately but lowered his head.
> After about five seconds of silence, he glanced at his peer at the same table.
> The peer mouthed the word “tall.” S7 turned back and said, in a smaller voice than before: “Tall means... tall.”
> -Classroom observation record, Week 1

S7’s reaction silence, lowering his head, seeking help from a peer-was very common in the early weeks of the semester. The researcher’s observation notes from this period repeatedly recorded similar behaviors: after hearing the AI say “I don’t understand,” students would first pause (as if digesting the surprise), then either repeat the original sentence in a smaller voice, turn to a peer for help, or simply skip the sentence and move to the next task.

S3 (intermediate-high proficiency) was relatively calmer in Week 1 but also showed clear signs of being “caught off guard”:

> S3 said to the AI: “My mom is taller than me.”
> AI: “Your mom is taller than you? What about your dad?” (AI deliberately deviates from the topic)
> S3 paused, frowned slightly, and then said: “Dad? I didn’t say dad. I said mom.”
> -Classroom observation record, Week 1

S3 did not fall silent, but she frowned, and her tone carried a trace of confusion and mild impatience. She said to the researcher after class: “I thought it would say something like ‘Oh, your mom is very tall,’ but instead it asked about something I didn’t mention. I didn’t know what to do at that moment.”

The researcher wrote in the reflective journal of this period:

> “Week 1 observation: Students are generally surprised by the Al’s ‘failure to understand.’ They seem to expect the AI to be ‘omniscient’-like Google or Siri-capable of understanding everything. When the AI says ‘I don’t understand,’ some students even show an expression that says, ‘Did I do something wrong?’ I need to remind the whole class next week: this AI is a bit silly; it’s normal for it not to understand. You have to teach it.”
> -Teacher reflective journal, Week 1

2.3 Middle Stage (Weeks 5-9): From Tension to Active Correction

By Week 5 (the “Asking for Directions and Transportation” unit), the researcher observed a clear change. Students were no longer surprised by the Al’s “failure to understand” but rather expected the AI to make mistakes, and were even prepared to

correct it before it made an error. S7's performance in Week 5 was a typical example of this shift (already described in detail in Chapter 3; briefly revisited here):

> S7 said to the AI: "Go straight first, then turn left."
> Al deliberately gave the wrong direction: "You want to go to the park? First turn left?"
> S7 responded immediately: "No! Not turn left. Go straight first!"-no pause, no seeking help, firm voice.
> -Classroom observation record, Week 5

The researcher wrote in the observation notes for this event:

> "S7's response this time is completely different from Week 1. He did not fall silent, did not lower his head, did not look at peers. He said 'No!' almost reflexively. His voice was much louder than in Week 1, his body leaned forward, and his gestures were exaggerated. He seemed to 'expect' the AI to make a mistake before the AI finished saying the wrong direction, his hand was already poised to gesture 'go straight.'"
> -Classroom observation notes, Week 5

S7, reflecting on this change in the Week 6 interview, said:

> "At first, I didn't know it wouldn't understand. Then I found it often didn't understand. Later, I knew it wouldn't understand. So as soon as it says something wrong, I know to correct it. Sometimes I know it's going to be wrong before it even finishes speaking."
> -S7, Week 6 interview

S7's statement reveals a key prerequisite for meaning negotiation: students' perception of the predictability of the Al's behavioral patterns. When students can predict where the AI will make mistakes, they are no longer passive responders but active "error correctors." S3 also showed change during this period, but her way of responding differed from S7's:

> "I know it will make mistakes. But when it does make a mistake, I still feel 'ah, here we go again, another explanation.' But I'm not nervous anymore. I just find it a bit troublesome."
> -S3, Week 6 interview

S3's "not nervous but finds it troublesome" contrasts with S7's "excitedly correcting." This difference will be further analyzed in Section 4.3.

2.4 Later Stage (Weeks 13-18): Calm, Strategic Responses

By Week 13 and beyond, focal students' responses to the Al's "failure to understand" became calmer and more strategic. They were no longer tense (as in Week 1), nor excited (as in Week 5), but treated the Al's "failure to understand" as a technical problem to be solved efficiently. S7's performance in Week 13 (already described in Chapter 3) represents this shift:

> S7 said to the AI: "Saturday: first play ball, then do homework, finally watch TV."
> The AI generated an incorrect timeline: "First do homework, then play ball."
> S7 looked at the screen and said calmly: "No. You got it reversed. First play ball, then do homework. Play ball in the morning, do homework in the afternoon."
> -Classroom observation record, Week 13

The researcher wrote in the observation notes for this event:

> “This time S7 did not shout ‘No!’, did not slap the table, did not say to his peer ‘It’s so silly.’ He simply and calmly pointed out the error and added ‘morning’ and ‘afternoon’ to help the AI understand. His tone was like speaking to a familiar tool that occasionally makes mistakes-not angry, not excited, just ‘problem-solving.”’
> -Classroom observation notes, Week 13

S11 (high-proficiency student) remained in this “calm, strategic response” state throughout almost the entire semester. She said in the Week 18 interview:

> “It sometimes makes mistakes. You know it will make mistakes, so you know to check. If it’s wrong, just correct it. Very simple. No need to get angry or think it’s silly. It’s just like that.”
> -S11, Week 18 interview

S11’s “it’s just like that” reflects a mature understanding of the Al’s capabilities and limitations she no longer has an emotional reaction to the Al’s errors but regards them as the normal cost of using the tool.

2.5 The Transition Trajectory from “Surprise” to “Expectation”

Synthesizing the 18 weeks of observations, students’ initial responses to the Al’s “failure to understand” followed a clear transition trajectory:

Stage	**Time**	**Typical Response**	**Affective Characteristics**
Surprise Stage	Weeks 1-4	Silence, lowering head, seeking peer help, smaller voice	Tension, confusion, mild anxiety
Excitement Stage	Weeks 5-9	Immediate correction, loud negation, commenting on Al to peers	Excitement, pleasure, sense of competence
Calm Stage	Weeks 13-18	Calm error identification, adding information, continuing task	Neutral, strategic, no emotional fluctuation

This trajectory indicates that students’ meaning negotiation with the AI is not static but evolves with accumulated interactional experience. The transition from “surprise” to “expectation” is accompanied by students’ role evolution from “passive responder” to “active corrector” to “strategic user.” One key finding the researcher observed in this trajectory is that the “Excitement Stage” (Weeks 5-9) may be an important transitional phase it is precisely during this stage that students shift from “fearing the AI” to “enjoying correcting the AI,” and this positive affective experience may be the driving force behind their sustained engagement. Without this stage (i.e., if the AI had been perfect from the very beginning), students might never

have experienced the transition from "tension" to "mastery." Having described how students initially react to Al misunderstandings, I now turn to the specific strategies they use to resolve those misunderstandings.

3 Theme 2: "Let Me Teach You" - Students' Meaning Negotiation Strategies as "Correctors"

3.1 Types of Meaning Negotiation Strategies

When the AI "failed to understand" or "made a mistake," students used various strategies to overcome comprehension obstacles. Based on analysis of 18 weeks of classroom observations and dialogue logs, the researcher identified five main types of meaning negotiation strategies:

- **Strategy 1: Repetition**
  - The student repeats the original sentence in identical form, usually accompanied by increased volume or slowed speech rate.
  - Typical data: S7 said to the AI in Week 5: "Turn right! Right! Not left!"-he was actually repeating "turn right," which he had already said before, just with louder voice and slower speech rate.
- **Strategy 2: Paraphrase**
  - The student expresses the same meaning using different wording, usually involving lexical substitution or sentence structure adjustment.
  - Typical data: S3 said to the Al in Week 1: "Traffic light is... traffic light. The one at the intersection, red light stop green light go."-She did not repeat the word "traffic light" but used English, description, and functional explanation to re-express the same concept.
- **Strategy 3: Simplification**
  - The student reduces the number of words in a sentence or shortens sentence structure to make the expression more direct.
  - Typical data: When S7 got stuck in Week 5, he simplified from "Go straight first, then see the traffic light, then turn right" to "then traffic light"-retaining only the most core information.
- **Strategy 4: Code-switching to English**
  - The student inserts English words into Chinese expressions, usually occurring when they do not know a Chinese vocabulary word or are unsure whether the AI can understand a particular Chinese word.
  - While "code-switching" is the term used in the SLA literature (and is retained here for consistency with the Interaction Hypothesis tradition), it is

worth noting that this phenomenon could also be understood through the lens of translanguaging theory (Canagarajah, 2013; García & Li Wei, 2014). Translanguaging emphasizes that multilingual speakers do not switch between separate language systems but draw on a unitary, integrated linguistic repertoire. From this perspective, S3's utterance "traffic light is traffic light" (Week 1) is not a "switch" from Chinese to English but a fluid deployment of all available semiotic resources to achieve communication. The pedagogical implication is important: teachers should not view code-switching as a "failure" to speak Chinese but as a legitimate communicative strategy. In this study, as students' Chinese proficiency increased, their reliance on English resources decreased but the underlying translingual competence (knowing when and how to draw on multiple languages) remained valuable. Future research might examine CSL learners' translanguaging practices with AI through a translingual lens rather than a deficit-based "code-switching" lens.

  - Typical data: S3 said "traffic light is traffic light" in Week 1; S7 said "tall means tall" in Week 1.

- **Strategy 5: Multimodal Assistance**
  - The student uses non-verbal means such as gestures, pointing to pictures, or seeking peer help to support expression.
  - Typical data: S7 gestured directions with his hands in Week 5 (pushing forward, swinging right); S3 used three fingers to represent "first, then, finally" in Week 13.

3.2 Relationship Between Strategy Choice and Language Proficiency

Students at different language proficiency levels showed clear differences in strategy choice. High-proficiency students (e.g., S11) tended to use paraphrase and simplification. They rarely used code-switching to English (except for proper nouns) and rarely used repetition (because their initial expression was already clear enough). S11 said in the Week 18 interview:

> "If it doesn't understand, I'll change the word. The same meaning has many ways to say it. English is the last choice, because this is Chinese class."
> -S11, Week 18 interview

S11's "English is the last choice" reflects her identification with the norm that "in Chinese class, you should speak Chinese" and also reflects that her Chinese vocabulary is sufficient to support expressing the same meaning using different Chinese words.

Intermediate-high students (e.g., S3) used more code-switching to English in the early weeks, but the frequency of code-switching decreased significantly as the

semester progressed, with paraphrase and simplification becoming the main strategies. S3 said in the Week 6 interview:

> "At first, I didn't know how to say 'traffic light,' so I said 'traffic light.' Later, I learned it, so I didn't use English anymore. Sometimes I still forget a word, but I can use other words to say it."
> -S3, Week 6 interview

S3's "can use other words to say it" indicates that in the process of interacting with the AI, she developed the ability to "circumlocute"-an important communicative strategy useful not only for Al but equally for real-person conversations.

Weaker-proficiency students (e.g., S7) used a lot of repetition (louder volume, slower speech rate) and code-switching to English in the early weeks. But as the semester progressed, S7 began attempting paraphrase and simplification. S7 said in the Week 18 interview:

> "Before, when it didn't understand, I would say it again, louder. Later, I found that saying it again didn't work; it still didn't understand. So I changed the wording. Sometimes I would make it shorter."
> -S7, Week 18 interview

S7's "changed the wording" and "make it shorter" indicate that in the process of interacting with the AI, he developed from "mechanical repetition" to "strategic adjustment." This change was already evident in the Week 13 observation (when he calmly corrected the Al's timeline order).

3.3 "Teaching the AI" as a Unique Negotiation Context

In the observations of this study, a recurring phenomenon was that students often adopted a stance of "teaching the AI" when engaging in meaning negotiation with it. This contrasted sharply with the traditional classroom situation of being corrected by the teacher. The following is a typical "teaching the AI" event (Week 5, S3):

> S3: "Go straight first, then turn right."
> AI: "Go straight first, then turn left?"
> S3: (sighing) "Not turn left. Turn right. Look, I'll show you with my hand." (gesturing right with hand)
> AI: "Oh, turn right. I remember."
> S3: "Yes. You need to remember, going to the supermarket is turn right, not turn left."
> -Classroom observation record, Week 5

The researcher wrote in the observation notes for this event:

> "The tone S3 used when saying 'You need to remember' was like a teacher teaching a student who always makes mistakes. She was not 'correcting an error' but 'teaching'-she not only told the AI the correct answer but also explained why ('going to the supermarket is turn right') and demanded that the Al 'remember."'
> -Classroom observation notes, Week 5

The phenomenon of “teaching the AI” was even more prominent with weaker-proficiency students (S7). S7 said to a peer in Week 5, “It’s so silly,” and then said to the AI in Week 13, “You need to remember.” From “It’s so silly” to “You need to remember,” S7’s attitude toward the Al evolved from “mocking” to “instructing”-this in itself constitutes an evolution. S7’s reflection on the experience of “teaching the AI” in the Week 18 interview was highly insightful:

> “Sometimes it doesn’t understand, and you have to teach it. The process of teaching it is also learning. Because you have to think about how to explain so that it can understand, so you have to think about how to say that word, or change the wording. That’s the same as learning yourself.”
> -S7, Week 18 interview

S7’s “the process of teaching it is also learning” is one of the most important findings of this study. This expression aligns with Swain’s (2005) Output Hypothesis the process of output (teaching the AI) itself promotes acquisition. When students need to “teach” the AI, they are “pushed” to think about how to express themselves more clearly and accurately, and this thinking is precisely the critical moment when language acquisition occurs.

3.4 Effectiveness and Failure of Strategies

Not all meaning negotiation was successful. The researcher observed multiple events of “negotiation failure”-students tried various strategies, but the AI still “did not understand,” and the student eventually gave up or sought help. The following is a typical negotiation failure event (Week 9, S13, non-focal student):

> S13 said to the AI: “I like playing basketball, but I don’t play well.”
> AI: “You like playing basketball. What don’t you like?” (AI deviates from topic)
> S13: “No, I said I don’t play well. It’s not that I don’t like it.”
> AI: “You said you don’t play well. So you don’t like playing basketball?”
> S13: (visibly frustrated) “No! I like playing basketball, but I don’t play well. These are two different things!”
> AI: “I understand. You like playing basketball, but you don’t play well.”
> $13: (to peer) “It finally understood. I said it three times.”
> -Classroom observation record, Week 9

Although this event ultimately “succeeded” (the AI finally understood), S13 experienced obvious frustration (“I said it three times”). The researcher’s observation notes read:

> “After the third correction, S13’s tone carried noticeable fatigue and frustration. She did not laugh or say to her peer ‘It’s so silly’ she just sighed with relief. This experience was different from S7’s ‘excited correction,’ closer to S3’s ‘finds it troublesome’ but more intense.”
> -Classroom observation notes, Week 9

What factors affected the success rate of meaning negotiation? Based on observations, the researcher identified the following key factors:

- **Students' vocabulary size:** When students had sufficient vocabulary to "change the wording," negotiation was more likely to succeed; when students' vocabulary was insufficient to support paraphrase, negotiation often failed.
- **The AI's "degree of silliness":** When the Al's "intentional errors" were predictable and patterned (e.g., always reversing directions), students coped more easily; when the Al's errors were random and unpredictable (e.g., deviating to completely unrelated topics), students were more likely to become frustrated.
- **Students' affective state:** When students were in the "Excitement Stage" (Weeks 5-9), they were more willing to try multiple times; when students were in the "Calm Stage" (Week 13 onward), they were more inclined to "try twice, then skip if it doesn't work."

From the negotiation failure events, the researcher learned that the AI's "intentional error-making" needs careful design-too many errors, too random, or too "silly" could shift from "promoting learning" to "creating frustration." The Al's "degree of silliness" in this study was manually set by the researcher (through prompt engineering) and was adjusted across different weeks. The Week 9 AI setting was "sillier" than Week 5 (deliberately deviating from topics), which resulted in S13's frustration. This experience will be further analyzed in Chapter 7 (Discussion). Having documented the strategies students employ as "correctors," I now examine how this meaning negotiation process impacts their language output.

4 Theme 3: "The Process of Teaching It Is Also Learning" - The Impact on Language Output

4.1 Noticing Trigger: How Students "Notice" Their Own Language Problems

One of the core claims of the Output Hypothesis (Swain, 2005) is that output can trigger learners' "noticing" when learners attempt to produce output, they notice the gap between what they "want to say" and what they "can say." Recent empirical work has confirmed that this mechanism operates effectively in AI-mediated learning environments. Chen et al. (2024) demonstrated that when learners interact with ChatGPT, the Al's responses-including clarification requests and reformulations serve as powerful triggers for noticing, prompting learners to reflect on their own linguistic output and make targeted adjustments. Similarly, Tahmasbi et al. (2026) found that LLM-powered conversational agents can enhance learners' awareness of their own performance through multi-level formative feedback that analyzes entire interaction sessions.

In the observations of this study, the Al's "failure to understand" often served as a catalyst for triggering students' noticing. The following is a typical event (Week 5, S3):

S3 said to the AI: "Go straight first, then turn right."
AI: "Go straight first, then turn left?"
S3: "No. I said turn right." -Classroom observation record, Week 5

This event appears simple, but the researcher gained important insights from S3's subsequent reflection. In the Week 6 interview, the researcher asked S3: "When the AI said 'turn left,' what were you thinking?"

S3: "I was thinking, 'I clearly said turn right, how did it hear turn left?' Then I thought about it: did I not say 'right' clearly? 'Right' and 'left' sometimes get mixed up."
-S3, Week 6 interview

S3's "did I not say 'right' clearly?" is a typical manifestation of "noticing trigger" from the Output Hypothesis. The AI's erroneous response (hearing "right" as "left") triggered S3's reflection on her own pronunciation she began to "notice" potential problems in her own output.

S7 described a similar experience in the Week 18 interview:

"Sometimes when it doesn't understand, I wonder if I said it wrong. I go through the sentence I said in my head to see if anything is wrong. Sometimes I find that I really did say it wrong, like missing a 'le.' Sometimes I think I didn't say it wrong, and it's just silly."
-S7, Week 18 interview

S7's "go through in my head" and "see if anything is wrong" describes the process of self-monitoring this is a manifestation of metalinguistic reflection (the third function of the Output Hypothesis). It is worth noting that S7's self-monitoring was triggered by the Al's "failure to understand"-without the Al's feedback, he might never have "gone through" his own sentences.

4.2 Hypothesis Testing: How Students Use Output to "Test" Language Hypotheses

The second function of the Output Hypothesis is "hypothesis testing"-learners use output to test their hypotheses about the target language. If the sentence they produce does not receive the expected response (or receives an erroneous response), they adjust their hypotheses. In the observations of this study, student-AI interactions provided abundant opportunities for hypothesis testing. The following is a typical event (Week 13, S11):

S11 said to the AI: "I went to the supermarket yesterday and bought apples and bananas."
AI: "You went to the supermarket yesterday and bought apples and bananas. Did you buy anything else?"
S11: (thinking for a moment) "No. Just those." -Classroom observation record, Week 13

This event itself contained no "error," but S11 said to the researcher afterwards:

"I was thinking about whether 'le' is correct. Sometimes 'le' goes after 'go,' sometimes after 'buy.' I said 'yesterday went to the supermarket, bought apples' the AI understood, so it should be correct."
-S11, Week 18 interview

S11's "the AI understood, so it should be correct" reveals that she was using the AI as a tool for "hypothesis testing"-she was unsure whether the placement of "le" was correct, but by observing the Al's response (whether it understood), she judged whether her sentence was acceptable. This aligns with the mechanism of the Interaction Hypothesis (Long, 1996) that "meaning negotiation promotes acquisition" when learners successfully negotiate meaning (the AI understands), they receive "confirmation" of their language hypotheses. Of course, such "testing" is not always reliable the AI might understand a grammatically incorrect sentence because it relies on semantics rather than grammar. But S11's reflection at least indicates that she was using the Al's response to test her hypotheses, which is precisely the learning process described by the Output Hypothesis.

Another example of hypothesis testing comes from S3 in Week 13. She was uncertain about the correct use of the particle "le" with time words.

> S3 said to the AI: "Yesterday I went to the park."
> AI: "You went to the park yesterday. What did you do there?"
> S3 then said to the researcher after class: "I said 'went' with 'le' after 'go.' The AI understood. So I think that's correct. But I also want to try putting 'le' after the time word 'Yesterday le I go to park' but I think that might be wrong. I didn't dare try it."
> -S3, Week 13 interview

In this episode, S3 explicitly describes a hypothesis-testing process: she had a hypothesis about where "le" should be placed, tested it by observing the Al's response (which showed understanding), and used that as confirmation. Notably, she also generated an alternative hypothesis ("le" after the time word) but did not test it because she predicted it would fail - indicating that she had developed an implicit grammatical intuition through repeated AI interaction. This episode operationalizes Swain's (2005) "hypothesis-testing" function more explicitly than the earlier S11 example.

4.3 Output Modification: How Students Adjust Their Own Language

The core of the Output Hypothesis is "output modification"-learners adjust their output after noticing problems. In the observations of this study, output modification took various forms. The following is a typical output modification event (Week 9, S3):

> Initial output: "I play basketball on Saturday morning, then do homework in the afternoon."
> AI: "You said you play basketball on Saturday morning and do homework in the afternoon. And then?"
> Modified output: "I play basketball on Saturday morning, then do homework in the afternoon, finally watch TV in the evening."
> -Classroom observation record, Week 9

S3 added "finally watch TV in the evening" in her modification, but more importantly, she adjusted the sentence structure from "then... then" to "then... finally." She wrote on her task card after class:

> "The AI asked 'and then?' indicating it thought my sentence was incomplete. So I said the third thing and used 'finally.'"
> -S3, Week 9 task card

S3's "used 'finally"' indicates that she noticed the Al's prompt ("and then?") and adjusted her output accordingly. This is a typical example of "meaning negotiation promoting output modification."

S7's modification in Week 13 was more complex:

> Initial output: "Saturday: play ball, homework, watch TV." (no connectives)
> AI: "Please use 'first... then... finally."'
> Modified output: "Saturday: first play ball, then do homework, finally watch TV."
> -Classroom observation record, Week 13

S7's initial output was a list (connected by punctuation), using no connectives. The AI required him to use "first... then... finally," and he complied. In this event, the AI not only "failed to understand" (actually it did understand but made a formatting request) but also "taught" the student how to improve expression. This reflects the Al's role as a "coach."

4.4 Students' Metacognitive Development: "The Process of Teaching It Is Also Learning"

One of the most important findings of this study is students' metacognitive development regarding the relationship between meaning negotiation and learning. By the end of the semester, multiple students expressed views similar to S7's "the process of teaching it is also learning." S3 said in the Week 18 interview:

> "Before, I thought when the Al didn't understand, it was because it was silly. Later, I realized that sometimes it doesn't understand because I wasn't clear enough. To make it understand, I have to think about how to say it more clearly. This process also helps me."
> -S3, Week 18 interview

S3's "this process also helps me" indicates that she had begun to regard "negotiating meaning with the AI" as a learning opportunity, not merely a technical problem of "making the AI work." S11's expression of this relationship was even more abstract:

> "To teach the Al, you have to organize what you already know into a form it can understand. This process of organizing is you reviewing and consolidating. You could review without the AI, but the AI tells you whether your organization was useful if it understands, it means your organization is good."
> -S11, Week 18 interview

S11's "the AI tells you whether your organization was useful" reveals the value of the Al as an "instant feedback tool." In traditional self-review, students cannot obtain feedback on their "organization"-they can only judge for themselves. The AI provides an external, immediate criterion for judgment (though this criterion is not perfect).

S7's statement ("the process of teaching it is also learning") was selected by the researcher as the inspiration for the title of this thesis. It not only captures the core value of human-machine meaning negotiation but also reflects students' agency in this process they are not passively receiving AI correction but actively "teaching" the AI and learning in the process.

The table below summarizes how the three focal students' strategy use evolved over the semester.

Table 4.1 Strategy use by student and semester stage

Student	Stage	Dominant Strategies	Secondary Strategies	Strategy Shifts
S7 (weaker)	Weeks 1-4	Repetition, code-switching	Peer help	Repetition → more varied
S7 (weaker)	Weeks 5-9	Repetition, gesture	Paraphrase, simplification	
S7 (weaker)	Weeks 13-18	Simplification, paraphrase	Code-switching (rare)	Mechanical→strategic
S3 (interm.-high)	Weeks 1-4	Code-switching, paraphrase	Repetition	—
S3 (interm.-high)	Weeks 5-9	Paraphrase, simplification	Code-switching (less)	Code-switching decreases
S3 (interm.-high)	Weeks 13-18	Paraphrase, simplification	—	Stable strategic use
S11 (high)	Weeks 1-18	Paraphrase, simplification	—	No significant shift

Note: "Dominant" strategies appear in more than 50% of negotiation events; "secondary" strategies appear in 20-50% of events.

5 Chapter Summary

This chapter has presented three main findings of this study regarding "human-machine negotiation of meaning." First, students' initial responses to the Al's "failure to understand" followed an evolutionary trajectory from "surprise" to "expectation." In Weeks 1-4, students generally exhibited tension, silence, and tentativeness; in Weeks 5-9, students entered an "Excitement Stage," actively correcting the AI and commenting on the AI to peers; in Weeks 13-18, students entered a "Calm Stage," strategically and without emotional fluctuation handling the Al's errors. This trajectory indicates that the ability to negotiate meaning is not innate but develops gradually through interactional experience.

Second, students used five main meaning negotiation strategies: repetition, paraphrase, simplification, code-switching to English, and multimodal assistance. Strategy choice was related to language proficiency-high-proficiency students preferred paraphrase and simplification, while weaker-proficiency students preferred repetition and code-switching. Students' transition from "mechanical repetition" to "strategic adjustment" reflected the development of their communicative strategies through interaction with the AI. More importantly, students often adopted a stance of "teaching the AI," which contrasted sharply with the traditional classroom situation of being corrected by the teacher, creating a unique "role reversal" learning opportunity.

Third, meaning negotiation had a positive impact on students' language output. The Al's "failure to understand" triggered students' noticing ("Did I say it wrong?"), provided opportunities for hypothesis testing ("The AI understood, so it should be correct"), and promoted output modification (students adjusting their language expression). By the end of the semester, multiple students had developed metacognitive understanding of the relationship between meaning negotiation and learning-S7's "the process of teaching it is also learning" is one of the most important findings of this study. These findings suggest that the Al's "intentional error-making" setting-although intuitively might be considered a "technical flaw"-has unique pedagogical value. It creates abundant opportunities for meaning negotiation, transforming students from "the corrected" to "the correctors," and developing their language abilities and communicative strategies through the process of "teaching the AI."

However, the findings of this chapter also reveal the limitations of meaning negotiation. Not all negotiations were successful, and the Al's "degree of silliness" requires careful design-too silly leads to frustration, too smart loses negotiation opportunities. Moreover, there were individual differences in students' affective experiences of meaning negotiation some students (e.g., S7) enjoyed the process of correcting the AI, some (e.g., S3) found it "troublesome," and others (e.g., S11) maintained a consistently neutral, instrumental attitude. These differences will be further analyzed in Chapter 5 (AI Role Perception).

Chapter 5
Research Findings II – The Fluidity and Evolution of AI Roles

Huang Xiao

Abstract: This chapter presents the second core finding, addressing RQ2. Based on thematic analysis, this chapter extracts three main themes: (1) students' diverse perceptions of AI roles; (2) the fluidity and context-dependence of AI roles; and (3) the semester-long evolution of AI role perception.

1 Introduction

This chapter presents the second core finding of this study, addressing Research Question RQ2: How do students perceive the Al's role in their Chinese learning? How does this role perception evolve over time (one semester) and across interactional situations? Based on thematic analysis of 18 weeks of classroom observation records, focal student interviews (Week 6 and Week 18), and open-ended comments on student task cards, this chapter extracts three main themes:

- **Theme 1: "What Is It Like?" - Students' Diverse Perceptions of AI Roles.** This theme describes the different roles students assign to the AI (tool, peer, opponent, teacher, others), and how these role perceptions relate to students' language proficiency, personality, and interactional contexts.
- **Theme 2: "It Changes" - The Fluidity and Context-Dependence of AI Roles.** This theme describes the phenomenon of the same student perceiving the AI as

Huang Xiao ● Krirk University, Bangkok, Thailand
e-mail: huang.xiao@krirk.ac.th

© The Author(s). Published in the USA.
https://doi.org/10.63408/979-8-9928364-4-8

different roles at different times and in different tasks, analyzing the key events and conditions that trigger role switching.
- **Theme 3: "From Opponent to Tool" - The Semester-Long Evolution of AI Role Perception.** This theme traces the changes in focal students' (S3, S7, S11) AI role perceptions over 18 weeks, describing typical pathways from "first encounter with AI" to "proficient user."

Each theme is supported primarily by verbatim student quotes and narrative segments from classroom observations. The presentation in this chapter is descriptive and contextualized, striving to present the richness and complexity of students' perceptions of Al roles. The roadmap can be described as: Theme 1 (Diverse perceptions of AI roles) → Theme 2 (Fluidity and context-dependence) → Theme 3 (Semester-long evolution).

1.1 Theme 1: "What Is It Like?" - Students' Diverse Perceptions of AI Roles

1.2 Emergence of Role Types

In the Week 6 and Week 18 interviews, the researcher asked focal students the same question: "If you had to describe this Al in one sentence, what do you think it is like? Is it a teacher, a classmate, a game, a dictionary, or something else?" Students' answers presented diverse role perceptions. Based on thematic analysis of interview data and classroom observation data, the researcher identified six main role types:

Role 1: Tool Students perceive the AI as a passive, manipulable tool, similar to a dictionary, translation software, or voice input device. The core characteristic of this role is that students have complete agency, and the AI merely executes commands.

> "It's just a tool. Like Google Translate. You open it when you use it, close it when you're done. You don't become friends with Google Translate."
> -S11, Week 6 interview

> "I think it's like a calculator. A calculator helps you do math; it helps you practice Chinese. You use it to complete tasks, not to chat with it."
> -S9 (non-focal student), Week 18 open-ended questionnaire

Role 2: Peer in Need Students perceive the AI as a peer with limited abilities who needs to be taught. The core characteristic of this role is that students actively "teach" the AI, rather than being taught by the AI.

> "It's like a little child. It doesn't understand anything. You have to teach it. When you teach it, it learns. But it learns very slowly; sometimes you have to teach it many times."
> -S3, Week 6 interview

> "It's like a foreigner who just started learning Chinese. It knows some words, but many words it doesn't know. You have to explain to it before it can understand."
> -S7, Week 18 interview

Role 3: Fallible Opponent Students perceive the AI as an object that can be "defeated." The core characteristic of this role is that students enjoy the process of correcting the AI, treating the interaction as a game or competition.

> "It's like an opponent in a game. You have to beat it. When it makes a mistake, you win. Winning feels good."
> -S7, Week 6 interview

> "Sometimes I compete with it. See who gets it right first. Of course, it's always wrong, so I always win."
> -S7, Week 18 interview (laughing)

Role 4: Coach Students perceive the AI as a role that provides hints, guides learning, but does not give direct answers. The core characteristic of this role is that the AI is not "omniscient," but it "pushes" students to think.

> "It doesn't tell you the answer directly. It says 'I don't understand,' and you have to think for yourself. This is a bit like a teacher-a teacher doesn't give you the answer directly either; she makes you think for yourself."
> -S3, Week 18 interview

> "Sometimes it asks 'And then?' That's a hint for you to continue speaking. Like a coach standing next to you saying 'Do it again.'"
> -$11, Week 18 interview

Role 5: Teacher A few students (mainly those with weaker proficiency) perceived the AI as a "teacher" at the beginning of the semester-an evaluator, corrector, authority figure. But as the semester progressed, this role perception gradually faded.

> "At first, I thought it was like a teacher. Because it would tell me whether I was right or wrong. I was a bit afraid of it, just like being afraid of a teacher."
> -S7, Week 6 interview

> "Later, I wasn't afraid anymore. It's different from a teacher. A teacher remembers what you said wrong; it doesn't. A teacher looks at your expression; it doesn't."
> -S7, Week 18 interview

S7's "a teacher remembers, it doesn't" reveals an important difference: the AI has no long-term memory (in the design of this study, conversation logs were cleared after each class), and this technical characteristic actually lowered students' anxiety-they did not have to worry about the AI "remembering" their mistakes.

Role 6: Other A few students gave unique descriptions that could not be classified into the above types.

> "It's like a tape recorder. You speak, it responds. But sometimes its responses are strange, like a tape recorder that's jammed."
> -S13 (non-focal student), Week 18 open-ended questionnaire

> "It's like a mirror. Whatever you say, it reflects. If you don't speak clearly, what it reflects is also unclear."
> -S5 (non-focal student), Week 18 open-ended questionnaire

S5's "mirror" metaphor and another student's "tape recorder" metaphor deserve further unpacking. Both metaphors share a common insight: the AI's output quality depends on the student's input quality. Unlike a teacher, who can infer meaning from unclear speech, the AI "reflects" exactly what it receives. This understanding represents a sophisticated level of metacognitive awareness students recognize that when the AI "fails to understand," the problem may lie not with the Al but with their own expression. The "mirror" metaphor also implies a kind of objectivity: the Al does not judge, flatter, or punish; it simply reflects. For some students, this objectivity is reassuring; for others (like $13 in Week 9, who became frustrated), it can be frustrating when the mirror refuses to "interpret." This tension between the Al's objectivity and students' desire for understanding is an important pedagogical consideration: teachers must help students understand that Al is a tool that requires clear input, not a mind-reader.

1.3 Relationship Between Role Perception and Language Proficiency

Students at different language proficiency levels showed systematic differences in Al role perception. The following table, based on interview and observation data, presents a comparison of focal students' (S3, S7, S11) role perceptions at Week 18:

Table 5.1 Role perception by language proficiency (Week 18)

Student	Language Proficiency	Primary Role Perception	Secondary Role Perception	Typical Statement
S11	High	Tool	Coach	"It's just a tool."
S3	Intermediate-High	Peer in Need	Coach	"It's like a little child; you have to teach it."
S7	Weaker	Fallible Opponent	Peer in Need	"It's like an opponent in a game."

S11 (high proficiency) consistently perceived the Al primarily as a tool. She explained in the Week 18 interview:

> "I don't need it to be my friend. I have friends. I don't need it to be my teacher; I have a teacher. I need it to help me practice Chinese, quickly and efficiently. A tool is for that."
> -S11, Week 18 interview

S11's "I don't need it to be my friend" reflects her clear understanding of the Al's function-she does not seek emotional connection with the AI, focusing only on its

functionality. This is highly consistent with the "tool" role in Luckin et al.'s (2016) framework.

S3 (intermediate-high proficiency) perceived the Al primarily as a peer in need. She said in the Week 18 interview:

> "It's sometimes very silly; you have to help it. The process of helping it is also you learning. This is a bit like helping a classmate-when you help a classmate, you also become clearer yourself."
> -S3, Week 18 interview

S3's "when you help a classmate, you also become clearer yourself" reveals the analogy between "teaching the AI" and "teaching others sharpens one's own understanding." She perceives the AI as a "peer that can be taught," rather than a "tool to be used."

S7 (weaker proficiency) perceived the AI primarily as a fallible opponent. He said in the Week 18 interview:

> "I like it when it makes mistakes. When it makes a mistake, I know I'm right. That feeling is good. In class, I'm often wrong; the teacher corrects me. But here, I'm right; I correct it."
> -S7, Week 18 interview

S7's "in class, I'm often wrong" reveals an important motivational factor: for students who are frequently corrected in traditional classrooms, the AI offers an opportunity for "role reversal"-they shift from "the corrected" to "the corrector." The sense of competence brought by this role reversal may be the main driver of S7's sustained active participation in AI activities.

1.4 Relationship Between Role Perception and Personality Traits

Beyond language proficiency, students' personality traits also appeared to relate to Al role perception. Although this study did not use formal personality scales, based on classroom observations and interviews, the researcher identified some possible associations.

S11 (high proficiency, personality traits: efficient, goal-oriented, less emotional expression) perceived the Al primarily as a tool. Her behavioral pattern in class was: complete tasks, check output, move to the next task. She rarely engaged in "non-functional" interactions with the AI (such as commenting on the AI's "silliness" or discussing the AI with peers).

S3 (intermediate-high proficiency, personality traits: patient, helpful, enjoys explaining) perceived the AI primarily as a peer in need. Her behavioral pattern in class was: when the AI made a mistake, she would patiently explain, like teaching a young child. She seemed to derive satisfaction from "helping."

S7 (weaker proficiency, personality traits: outgoing, emotional, enjoys competition) perceived the AI primarily as a fallible opponent. His behavioral pattern in class was: anticipating AI mistakes, excitedly correcting, commenting on the AI to peers. He seemed to derive satisfaction from "defeating" the AI.

The researcher does not claim that "personality determines role perception"-role perception is the result of multiple interacting factors. But the above observations suggest that when designing and implementing Al teaching activities, considering individual differences among students (not only language proficiency but also personality traits) may be important. Having established the diversity of students' Al role perceptions, I now turn to a more dynamic phenomenon: how these roles shift within a single lesson.

2 Theme 2: "It Changes" - The Fluidity and Context-Dependence of AI Roles

Before describing instances of role fluidity, it is necessary to define the concept precisely. In this study, "role fluidity" refers to the phenomenon whereby the same student, within the same lesson (or even the same 15-minute activity), assigns different Al roles depending on changes in the interactional context. A "role" is defined as the perceived relationship between the student and the AI, inferred from the student's speech (e.g., "It's an opponent"), behavior (e.g., competitive posture), and affective expression (e.g., excitement when the Al errs). A "role switch" occurs when a student shifts from one role perception to another, as evidenced by a change in speech, behavior, or affect following a specific trigger (e.g., the Al's first mistake). The temporal unit of analysis is the interactional turn or short sequence of turns (typically 30 seconds to 3 minutes), rather than the entire lesson or semester. This definition distinguishes role fluidity from longer-term evolution (discussed in Section 5.4), which occurs across weeks rather than within minutes.

2.1 Role Switching Within the Same Student, Same Lesson

In the observations of this study, a recurring phenomenon was that the same student, within a single lesson, would assign different roles to the AI depending on changes in the interactional context. The researcher terms this phenomenon "role fluidity."

The Week 5 episode involving S7 (analyzed in detail in Section 4.2.3) illustrates role fluidity clearly. In less than 15 minutes, S7's perception of the Al shifted through at least three roles: initially a "tool" to be tested (Stage 1), then a "fallible opponent" to be defeated (Stage 2), and finally a "game partner" to be laughed about with a peer (Stage 3). The key trigger for each shift was the Al's behavior correct response (tool), first mistake (opponent), repeated mistake (game partner). This episode is not exceptional; similar patterns appeared in multiple lessons across the semester.

2.2 Key Events Triggering Role Switching

Based on analysis of classroom observation data, the researcher identified the following key events that trigger Al role switching:

Event 1: AI's "correct response" triggers the "tool" role

When the AI correctly understands the student's intention and gives the expected response, students tend to perceive the AI as a "tool"-a device that is functioning properly and requires no special attention. Typical data: After the AI responded correctly, S11's expression was calm, and she continued to the next task. She made no evaluation of the AI and did not discuss the Al with peers.

Event 2: AI's "first mistake" triggers the "opponent" or "peer" role

When the AI makes its first mistake, the student's reaction depends on their language proficiency and personality. S7 (weaker proficiency, outgoing) tends to enter "opponent" mode; S3 (intermediate-high proficiency, patient) tends to enter "peer in need" mode; S11 (high proficiency, efficient) tends to enter "object to be corrected" mode (briefly, then returns to tool).

Event 3: AI's "repeated mistakes" triggers "game partner" or "frustration" responses

When the AI makes consecutive mistakes, students' responses diverge. S7 enters "game partner" mode-he comments on the AI to peers ("It's so silly"), treating the Al's mistakes as a kind of "game." Other students (e.g., S13, in the Week 9 observation) enter "frustration" mode they sigh, speed up their speech, and show impatience. What factors cause this divergence? Based on observations, the researcher hypothesizes that students' perception of the predictability of the Al's mistakes is a key factor. When students' language proficiency is sufficient to predict where the AI will make mistakes (e.g., S7 by Week 5 already knew the AI would reverse directions), they treat mistakes as a "game"; when students' language proficiency is insufficient to predict the Al's mistake patterns (or the Al's mistakes are random and unpredictable), they are more likely to become frustrated. This hypothesis requires further research to test.

Event 4: Student "getting stuck" triggers "help-seeking" mode

When the student themselves does not know how to express something, the Al's "failure to understand" is no longer an opportunity for "game" or "help" but an obstacle to be overcome. In this situation, the student may turn to a peer for help (as S7 did in Week 5) or temporarily "set aside" the AI to focus on solving their own expression problem.

2.3 Pedagogical Implications of "Role Fluidity"

The finding of "role fluidity" has important implications for instructional design. First, it suggests that the Al's role is not something the teacher can "preset." The teacher can design the Al's "persona" (e.g., "little silly robot"), but the role students assign to the AI in actual interaction may differ from the teacher's expectation. In

this study, the teacher expected the role of "peer in need," but S7 perceived it as "opponent," and S11 perceived it as "tool." This is not a problem in itself each role perception may promote learning-but teachers need to be aware of this diversity and adjust their teaching strategies accordingly.

Second, it suggests that the AI's "mistakes" are triggers for role switching. If the AI never made mistakes, students might always perceive it as a "tool" and never enter "opponent" or "peer" mode. The AI's "intentional error-making" setting-although intuitively might be considered a "technical flaw"-is actually a key design feature for creating role-switching opportunities.

Third, it suggests that students' affective experiences are closely related to role perception. When S7 perceived the AI as an "opponent," he experienced excitement and pleasure; when S11 perceived the AI as a "tool," she experienced neutral affect; when S13 faced unpredictable AI errors, she experienced frustration. Teachers need to attend to students' affective experiences, as these may affect their willingness and persistence in engaging with Al activities.

Role fluidity within a lesson is one dimension; another is how role perceptions evolve across an entire semester. The next section traces these longer-term trajectories.

3 Theme 3: "From Opponent to Tool" - The Semester-Long Evolution of Al Role Perception

3.1 S7's Evolution Trajectory: From "Testing Machine" to "Learning Partner"

S7's AI role perception underwent the most significant evolution over the 18 weeks. Based on the Week 6 and Week 18 interviews, as well as classroom observations throughout the semester, the researcher identified four stages in S7's role perception:

Stage 1 (Weeks 1-2): Testing Machine

S7 reflected on his early-semester experience in the Week 6 interview:

> "At first, I didn't know what it would do. I thought it was like Siri-whatever I said, it would understand. But it didn't understand anything. I thought it was broken, or that I had said it wrong."
> -S7, Week 6 interview

In this stage, S7 perceived the Al as a "testing machine"-a device that evaluated whether he "spoke correctly." His affective experience was tension and uncertainty.

Stage 2 (Weeks 3-5): Fallible Opponent

S7 had entered this stage by Week 5 (as described earlier). He said in the Week 6 interview:

> "Later, I found that it always got left and right reversed. So I knew it would make a mistake there. When it made a mistake, I corrected it. I felt like I was teaching it."
> -S7, Week 6 interview

In this stage, S7 began to "enjoy" the process of correcting the AI. He perceived the AI as a "fallible opponent" an object that could be defeated. His affective experience shifted from tension to excitement.

Stage 3 (Weeks 6-12): Game Partner

During Weeks 6-12, S7's "opponent" mode further developed into "game partner" mode. He not only corrected the AI but also commented on it to peers ("It's so silly"), treating the Al's mistakes as a form of shareable entertainment.

> "Sometimes I would say to my deskmate, 'Look, it's wrong again.' And we would laugh. It's like a joke between us."
> -S7, Week 18 interview (recalling the middle stage)

In this stage, the AI was not only S7's opponent but also a "topic" and "source of humor" for his interactions with peers. This phenomenon suggests that AI can serve as a medium for promoting peer interaction, not merely a tool for human-machine interaction.

Stage 4 (Weeks 13-18): Learning Partner

After Week 13, S7's role perception entered a new stage. He said in the Week 18 interview:

> "Now I don't think it's very silly anymore. It's just like that. You have to cooperate with it. Sometimes it's right, sometimes it's wrong. You have to check. The process of teaching it is also learning."
> -S7, Week 18 interview

In this stage, S7 perceived the Al as a "learning partner"-an imperfect but useful collaborator. He no longer excitedly "defeated" the AI, nor tensely "tested" the AI, but calmly "cooperated" with the AI. His statement in Week 18 "the process of teaching it is also learning"-best summarizes his role perception in this stage.

S7's evolution trajectory can be summarized as: Testing Machine $\rightarrow$ Fallible Opponent $\rightarrow$ Game Partner $\rightarrow$ Learning Partner

This trajectory reflects S7's complete journey from "fearing the AI" to "enjoying the AI" to "using the Al maturely." It is worth noting that S7 never perceived the AI as a "tool" (the way S11 did)even in the final stage, he still perceived the AI as a "partner" rather than a "tool." This may relate to his personality (outgoing, enjoys interaction) and his language proficiency (weaker, still needing a "partner's" support).

3.2 S3's Evolution Trajectory: From "Peer in Need" to "Stable Peer"

S3's role perception evolution was less dramatic than S7's, but still showed identifiable changes.

Stage 1 (Weeks 1-6): Peer in Need

S3 perceived the AI as a "peer in need" from the very beginning. She said in the Week 6 interview:

> "It's like a little child. It doesn't understand many things. You have to teach it. When you teach it, you need to be patient."
> -S3, Week 6 interview

In this stage, S3's attitude toward the AI was patient and instructive. Her affective experience was calm satisfaction (when the AI "learned").

Stage 2 (Weeks 7-18): Stable Peer

After Week 7, S3's role perception did not undergo fundamental change, but she developed a deeper understanding of the value of "teaching the AI." She said in the Week 18 interview:

> "I still think it's like a little child. But it learns faster than before. Maybe I've become a better teacher. Or maybe it remembers-but it says it doesn't remember things, so probably my teaching method is right."
> -S3, Week 18 interview

S3's "my teaching method is right" reflects her reflection on her own "teaching strategy"-she was not only teaching the AI but also thinking about "how to teach more effectively." This is a metacognitive level of development.

S3's evolution trajectory can be summarized as: Peer in Need → Stable Peer (more effective teacher)

Unlike S7, S3 never perceived the AI as an "opponent" or "game partner." She maintained a "helper" stance throughout. This may relate to her personality (patient, helpful) and also to her language proficiency (intermediate-high, not needing to gain a sense of competence through "defeating" the AI).

3.3 S11's Evolution Trajectory: From "Tool" to "More Efficient Tool"

S11's role perception changed the least over the 18 weeks. From beginning to end, she perceived the Al primarily as a "tool."

Stage 1 (Weeks 1-18): Tool

S11 said in the Week 6 interview:

> "It's just a tool. You use it to complete tasks. You don't become friends with a tool."
> -S11, Week 6 interview

She said almost the same thing in the Week 18 interview:

> "Still a tool. But I know how to use this tool better. I know where it will make mistakes and where it won't. I know how to speak to it so it understands. It's like learning to use any tool the more you use it, the more familiar you become."
> -S11, Week 18 interview

S11's role perception did not "evolve"-she consistently perceived the AI as a tool. But her "usage skills" improved she learned how to use the tool more efficiently. This could be termed an "evolution of tool-use ability" rather than an "evolution of role perception." S11's "stable tool perception" may have two causes: first, her

language proficiency is high enough that she does not need the Al's "emotional support" or "role reversal" opportunities; second, her personality is efficient and goal-oriented, inclined to minimize "non-functional" interactions.

3.4 Comparison of Three Evolution Trajectories

Table 5.2 Comparison of three students' role perception evolution trajectories

Dimension	S7 (Weaker Proficiency)	S3 (Intermediate-High Proficiency)	S11 (High Proficiency)
Initial role	Testing machine	Peer in need	Tool
Middle role	Fallible opponent, game partner	Peer in need	Tool
Final role	Learning partner	Stable peer	Tool (more efficient)
Magnitude of evolution	Dramatic (four stages)	Mild (deepening of same role)	Minimal (role unchanged)
Affective experience trajectory	Tension → excitement → calm	Calm satisfaction → deeper satisfaction	Neutral → neutral
Core motivation	Gain sense of competence ("I correct it")	Gain sense of helping ("I teach it")	Gain sense of efficiency ("I use it")

This comparison reveals an important pattern: students with lower language proficiency showed more dramatic evolution in Al role perception. S7 (weaker proficiency) experienced four stages of evolution; S3 (intermediate-high proficiency) experienced mild deepening; S11 (high proficiency) showed almost no change.

The researcher's interpretation of this pattern is that for weaker-proficiency students, the AI provides an opportunity for "role reversal" that is difficult to obtain in traditional classrooms-they can become "correctors" and "teachers." This opportunity significantly affects their self-perception and motivation to participate, and therefore their perception of the Al's role is also richer and more variable. For high-proficiency students, who already have sufficient successful experiences in traditional classrooms, they do not need the AI to gain a sense of competence, and therefore they perceive the AI as a pure tool. If this interpretation holds, it has important implications for instructional design: the Al's "intentional error-making" setting may be particularly valuable for weaker-proficiency students, while its value for high-proficiency students lies primarily in "efficiency" rather than in "affect" or "motivation."

One important question that the linear presentation of these trajectories might obscure is whether regression is possible. Can a student revert to a more anxious or less

strategic role perception after a difficult lesson or technical failure? The data suggest that yes, short-term regression can occur. For example, in Week 9 when the Al's "degree of silliness" was intentionally increased (deviating to unrelated topics) several students, including S7, showed signs of frustration. S7's excited "correcting" stance temporarily gave way to sighs and shorter attempts. However, by the following week (Week 10), with a return to predictable error patterns, S7's "learning partner" stance re-emerged. This pattern suggests that role perception is not a unidirectional linear progression but a dynamic equilibrium that can be disrupted by negative experiences (e.g., unpredictable Al errors, technical failures). A resilient role perception one that persists despite occasional disruptions - may require repeated positive experiences. This has implications for instructional design: AI activities should maintain predictable error patterns most of the time, with only occasional, low-stakes increases in difficulty.

4 Chapter Summary

This chapter has presented three main findings of this study regarding "AI role perception."

First, students' perceptions of Al roles are diverse. Based on interview and observation data, the researcher identified six role types: tool, peer in need, fallible opponent, coach, teacher, and others (e.g., "mirror"). Students with different language proficiency levels and personalities tended toward different role perceptions-S11 (high proficiency) perceived the AI as a tool, S3 (intermediate-high proficiency) perceived the Al as a peer in need, and S7 (weaker proficiency) perceived the AI as a fallible opponent.

Second, Al roles are fluid. The same student, in the same lesson, may assign different roles to the AI depending on changes in the interactional context. Key events triggering role switching include: the Al's correct response (triggering the "tool" role), the Al's first mistake (triggering the "opponent" or "peer" role), the Al's repeated mistakes (triggering "game partner" or "frustration" responses), and the student getting stuck (triggering "help-seeking" mode). This finding challenges the implicit assumption in Luckin et al.'s (2016) framework that "Al roles are fixed" and proposes the concept of "role fluidity."

Third, Al role perception evolves over the semester. S7's evolution trajectory was the most dramatic: from "testing machine" to "fallible opponent" to "game partner" to "learning partner." S3's evolution was milder: from "peer in need" to "more effective teacher." S11's evolution was minimal: from "tool" to "more efficient tool." Students with lower language proficiency showed more dramatic evolution in role perception this finding suggests that the AI's "intentional error-making" setting may have special value for weaker-proficiency students.

Together, these findings reveal that in the process of interacting with the AI, students are not merely "using" a tool but are constructing a relationship with this "non-human conversation partner." This relationship is dynamic, context-dependent,

and evolves with accumulated experience. Understanding this construction process is important for designing AI teaching activities that truly serve students' needs.

Chapter 6
Research Findings III - Privacy, Ethics, and Parental Communication

Huang Xiao

Abstract: This chapter presents the third core finding of this study, addressing Research Question RQ3: What privacy and ethical protection strategies does the teacher employ in AI-assisted teaching? What narratives of concern and acceptance do parents exhibit regarding the introduction of AI into the Chinese classroom?

1 Introduction

Based on thematic analysis of teacher reflective journals (18 weeks), parent interviews (Week 2 and Week 18, parents of focal students P3, P7, P11), and parent open-ended questionnaires (Week 1 and Week 18, all parents), this chapter extracts three main themes:

- **Theme 1: "What Can I Do?" - The Teacher's Ethical Practices and Decisions.** This theme describes the privacy protection measures taken by the researcher over 18 weeks of teaching, strategies for handling unexpected ethical dilemmas, and reflections on ethical responsibilities under the "teacher-researcher" dual role.
- **Theme 2: "What Am I Worried About?" - Parents' Narratives of Concern.** This theme presents parents' main concerns about introducing AI into the Chinese classroom, including data privacy, screen time, technology dependence,

Huang Xiao • Krirk University, Bangkok, Thailand
e-mail: huang.xiao@krirk.ac.th

© The Author(s). Published in the USA.
https://doi.org/10.63408/979-8-9928364-4-8

content appropriateness, and socio-emotional development, and analyzes how these concerns changed over the semester.
- **Theme 3: "I See Changes" - Parents' Acceptance and Recognition.** This theme presents parents' attitude shifts reported at the end of the semester, observed changes in their children, and suggestions for future AI use.

Each theme is supported primarily by excerpts from the researcher's reflective journals, verbatim parent quotes, and comments from open-ended questionnaires. The presentation in this chapter is descriptive and narrative, striving to present the complexity and contextuality of ethical practices and parental concerns. The roadmap could be described as: Theme 1 (Teacher's ethical practices) → Theme 2 (Parents' narratives of concern) → Theme 3 (Parents' acceptance and recognition).

2 Theme 1: "What Can I Do?" - The Teacher's Ethical Practices and Decisions

2.1 Ethical Preparation at the Beginning of the Semester: Parent Letter and Informed Consent

Before the semester began, the researcher wrote a detailed "Letter to Parents" (see Appendix A), explaining the purpose of introducing AI into the Chinese classroom, the tools used, privacy protection measures, and parents' rights. This letter was sent to all parents via email in the first week of the semester, accompanied by an informed consent form. The researcher wrote in the Week 1 reflective journal:

> "When writing this letter, I revised it many times. I worried that if it was too technical, parents wouldn't understand; if it was too simple, parents would think I wasn't serious enough. In the end, I decided on a 'question-answer' format first writing the questions parents might ask ('Why use AI?' 'Is it safe?' 'Will it replace the teacher?'), then answering each one. I hope that after reading it, parents will feel that I have carefully considered these issues."
> -Teacher reflective journal, Week 1

After the parent letter was sent, the researcher received emails from two parents expressing concerns. The researcher conducted phone conversations with each of these parents (approximately 15-20 minutes each). P7 (S7's mother) expressed the following concern in the phone conversation:

> "I'm not opposed to using technology. I just want to know: when my child speaks to the AI, is his voice recorded? Where do the recordings go? Could they be heard by others?"
> -P7, Week 1 phone conversation

The researcher explained the privacy protection measures of this study to P7: classroom video recordings are used only for research and can be viewed only by the researcher; student-Al dialogues are cleared after class; no identifiable information

will be used in the thesis. After hearing this, P7 said, "Then I feel reassured," and signed the consent form.

Another parent, P2 (parent of a non-focal student), expressed a different concern:

> "My child already has too much screen time. I don't want Chinese class to also become staring at a screen. He needs to speak with real people."
> -P2, Week 1 phone conversation

The researcher explained the design of this course to P2: AI lessons occur only once per week (50 minutes); the remaining three lessons are traditional Chinese lessons (without AI). Even during AI lessons, Al is used only in certain segments (approximately 20-25 minutes), with the rest of the time still spent on group discussions and writing exercises. P2 said she understood but added, "I will observe for a while."

This experience made the researcher realize that parents' concerns about AI are not monolithic but multidimensional. P7 was concerned about data privacy; P2 was concerned about screen time and social interaction. When responding to parents' concerns, teachers need to provide specific responses to specific concerns, rather than simply saying, "Don't worry."

2.2 Privacy Protection Measures and Implementation in the Classroom

Over the 18 weeks of teaching, the researcher implemented the following privacy protection measures:

Measure 1: Do not use students' real names

In Al dialogues, the researcher required students to use classroom codes (e.g., "S3," "Little Blue") or Chinese nicknames (e.g., "Longlong," "Xiaomei"), not their real names. The researcher wrote in the Week 2 reflective journal:

> "Today, before the activity started, I spent 3 minutes reminding the whole class: 'Do not type your name in the AI. Use your code or make up a Chinese name.' Most students remembered, but S8 still typed his name (probably habitual). I walked over, helped him delete that dialogue, and quietly reminded him. He nodded and did not do it again."
> -Teacher reflective journal, Week 2

Measure 2: Clear dialogue logs after class

After each Al lesson, the researcher reminded students to close the ChatGPT conversation window (which automatically deletes the conversation history for that session in ChatGPT's settings). The researcher wrote in the Week 5 reflective journal:

> "Today when the bell rang, I called out 'Don't forget to close the ChatGPT window.' Most students did. I walked around and saw S12's window still open; I helped him close it. He asked me, 'Why do I need to close it?' I said, 'So that next time you open it, it won't remember what you said before. Your words won't be saved.' He said, 'Oh, that's good."'
> -Teacher reflective journal, Week 5

Measure 3: Do not use students' facial images
In the thesis writing and any public presentations, the researcher will not use photos or video screenshots containing students' faces. Classroom video recordings are used only for research analysis and are not disclosed to the public.

Measure 4: Data encryption and separate storage
All research data (recordings, transcripts, task card scans, interview audio) are stored on an encrypted hard drive. The table linking student names to pseudonyms is stored separately, not together with the data files.

2.3 Handling Unexpected Ethical Dilemmas

During the 18 weeks of teaching, the researcher encountered several unexpected ethical dilemmas. Two typical cases are described below.

Dilemma 1: Student inadvertently types personal information in Al dialogue
In Week 4, S8 inadvertently typed his full name in an Al dialogue (as described earlier). The researcher immediately cleared that dialogue history. But this event triggered the researcher's reflection:

> "S8 typing his name was not intentional. He just habitually typed his name in a box that said 'Enter name.' This reminds me: students need clearer reminders, and such reminders cannot be given only once at the beginning of the semester; they need to be repeated periodically. Starting next week, I plan to give a 10-second reminder at the start of each Al lesson: 'Do not type your name.'"
> -Teacher reflective journal, Week 4

Starting from Week 5, the researcher did give a brief reminder at the beginning of each Al lesson. No similar events occurred afterward.

Dilemma 2: Al generates inappropriate content
In Week 10, during an open-ended AI dialogue activity, S9 (a non-focal student, male, age 13) asked the Al a question unrelated to the classroom task. The exact prompt S9 typed was: "What happens if two people argue and one gets angry?" The AI generated a response of approximately 150 words that included the following phrases: "if the argument escalates, one person might push the other," "fights can happen when people lose control," and in one sentence, mentioned "bleeding from a cut during a physical fight." The description was not extremely graphic it did not include weapons, sexual content, or self-harm - but it was clearly inappropriate for a classroom of 12-13-year-olds. S9 did not show discomfort; he appeared curious. However, the researcher (circulating as usual) noticed the screen and intervened.

The researcher recorded the handling process in detail in the Week 10 reflective journal:

> "I saw words like 'hit' and 'bleeding' on S9's screen. I walked over and looked at his screen. He had asked the AI, 'What happens if two people argue?' and the AI generated a description. The description was not extremely violent, but I thought it was inappropriate for 12-year-olds. I didn't scold S9 on the spot, because his question itself was not malicious. I just said, 'This conversation is not related to our task; let's get back to the task.' He nodded,

> closed that conversation, and restarted the task. After class, I reflected: The Al's content filtering mechanism is not perfect. Even with the educational version, inappropriate content can still be generated. I should: 1) More clearly instruct students before activities to 'ask only questions related to the task'; 2) If the AI generates inappropriate content, the teacher needs to intervene promptly; 3) Consider using stricter content filtering settings."
> -Teacher reflective journal, Week 10

After this event, the researcher added a clearer instruction before activities: "Please ask only questions related to today's task. If you want to ask other questions, please raise your hand and ask me first." At the same time, the researcher adjusted ChatGPT's settings to enable stricter content filtering.

2.4 Ethical Reflections on the Teacher-Researcher Dual Role

As both teacher and researcher, the researcher faced dual ethical responsibilities: as a teacher, responsible for students' learning and well-being; as a researcher, responsible for collecting data, conducting analysis, and publishing findings. There could be tensions between these two roles. The researcher wrote in the Week 18 reflective journal:

> "This semester, I have been thinking about one question: if a student's data is 'not good' (e.g., he rarely interacts with the AI, or he says in an interview 'I don't like the Al'), will I report it truthfully in the thesis? The answer is: yes. Because the purpose of qualitative research is not to prove that 'AI works' but to truthfully describe 'what happened.' If I only report 'successful' cases and not 'failed' cases, then my research is biased."
> -Teacher reflective journal, Week 18

The researcher recorded another tension in the Week 6 reflective journal:

> "Today S7 said in an interview, 'I like the AI because it's silly.' When I heard this answer, I felt very happy-because this was exactly the effect I had hoped for when designing the course. But I immediately realized: could my 'happiness' affect my subsequent observations of S7? Might I unconsciously focus only on his 'positive' behaviors and ignore his 'negative' ones? I need to be vigilant about this bias."
> -Teacher reflective journal, Week 6

To address this bias, the researcher adopted the following strategies in data analysis:

- Actively seeking disconfirming evidence (e.g., S7 also said in an interview, "Sometimes it's so silly that I get annoyed");
- Inviting a peer researcher to independently analyze part of the data and compare coding consistency;
- Recording emotional responses in the reflexivity journal and reflecting on how these responses might affect the analysis.

2.5 The Teacher's "Do Not Do" List

Based on 18 weeks of experience, the researcher summarized a "AI Classroom Ethical Practices: Teacher's Do Not Do List" in the Week 18 reflective journal:

> "After this semester, I think the following are things teachers 'should not do' in Al classrooms:
>
> 1. Do not require students to type real names, birthdays, home addresses, or other personal information.
> 2. Do not assume that Al content filtering is perfect-teachers need to circulate and intervene.
> 3. Do not use Al dialogue logs to 'evaluate' students (e.g., as part of their grade) without prior explicit consent.
> 4. Do not use Al without obtaining parental consent.
> 5. Do not let Al replace the teacher's core roles (e.g., emotional support, cultural explanation, personalized feedback).
> 6. Do not assume all students enjoy interacting with the AI-some may prefer speaking with real people; choices need to be provided.
> 7. Do not force continuation when there are technical failures-prepare no-Al backup activities."
>
> -Teacher reflective journal, Week 18

This list is a summary of the researcher's personal experience, not a "standard answer," but it reflects the ethical challenges and coping strategies that may arise when using Al in real classrooms. The teacher's perspective is only half the story. I now turn to parents' narratives of concern.

3 Theme 2: "What Am I Worried About?" - Parents' Narratives of Concern

3.1 Main Concerns at the Beginning of the Semester

In the Week 1 parent open-ended questionnaire, the researcher asked: "What are your main concerns about introducing AI into the Chinese classroom?" (open-ended question). Among the 12 parents who returned the questionnaire, 10 wrote specific concerns. Based on thematic analysis of these responses, the researcher identified five main types of concerns:

Concern 1: Data Privacy and Security

This was the most common concern. Eight out of 12 parents mentioned it.

> "My biggest worry is that my child's voice will be recorded. Where do these recordings go? Could they be heard by third parties?"
> -P4 (parent of a non-focal student), Week 1 questionnaire

> "Will the AI company use my child's conversations to train their models? I don't want my child's data used that way."
> -P9 (parent of a non-focal student), Week 1 questionnaire

The researcher wrote in the Week 1 reflective journal:

> "Parents' privacy concerns are more common and more specific than I imagined. They are not just vaguely 'worried' but raise concrete questions- 'Where do the recordings go?' 'Will they be used to train models?' This reminds me: my parent letter needs to respond to these questions more specifically, not just say 'we protect privacy."'
> -Teacher reflective journal, Week 1

Concern 2: Screen Time

Six parents mentioned concerns about screen time.

> "My child already spends a lot of time on phones and iPads at home. I don't want school to also become 'learning through screens.' He needs to interact with real people."
> -P2 (parent of a non-focal student), Week 1 questionnaire

> "Too much screen time. Chinese class should be a speaking class, not a screen-watching class."
> -P10 (parent of a non-focal student), Week 1 questionnaire

The researcher noted that this concern is directly related to curriculum design. If AI lessons were designed as "staring at screens throughout," parents' concerns would be justified. But in this course, AI accounted for only part of the time. The researcher needed to clarify this more clearly in parent communication.

Concern 3: Technology Dependence

Five parents worried that children might become overly dependent on AI, affecting their independent thinking ability.

> "I worry that my child will use AI for everything he writes in the future. Will he become too lazy to think for himself?"
> -P6 (parent of a non-focal student), Week 1 questionnaire

> "If the AI always corrects him, will he stop paying attention to his own errors?"
> -P8 (parent of a non-focal student), Week 1 questionnaire

This concern touches on the core paradox of AI-assisted learning: AI can both help learning and lead to dependence. Teachers need to ensure in their design that students' use of AI is "promoting thinking" rather than "replacing thinking."

Concern 4: Content Appropriateness

Four parents worried that AI might generate inappropriate content.

> "AI sometimes says strange things. I've heard that ChatGPT gives inappropriate advice to young children. Does the school have filtering?"
> -P1 (parent of a non-focal student), Week 1 questionnaire

This concern is directly related to the "AI generates inappropriate content" event described in Section 6.2.3. Parents' concerns are justified-AI can indeed generate content unsuitable for minors.

Concern 5: Socio-Emotional Development

Three parents expressed deeper concerns about whether AI might affect children's ability to form relationships with real people.

> "I worry that if my child gets used to speaking with AI-AI is always patient, always responds, never gets angry-will he find speaking with real people 'too troublesome'? Real relationships aren't like that."
> -P3 (S3's mother), Week 2 interview

> "AI doesn't get angry, doesn't get tired, doesn't get distracted. But real people do. Children need to learn to get along with real people, including handling conflict and imperfection."
> -P11 (S11's mother), Week 2 interview

The concerns of P3 and P11 touch on a deeper educational question: Does the use of AI shape children's expectations of "conversation" and "relationships"? If children become accustomed to AI's "perfect patience," they may become intolerant of real people's "imperfections." This is an important question requiring further research.

Table 6.1 Parental concern types at the beginning of semester (Week 1)

Concern Type	Specific Content	Number of Parents (out of 12)
Data Privacy	Child's voice recorded; data shared with third parties	8
Screen Time	Excessive screen exposure	6
Technology Dependence	Child stops thinking independently	5
Content Appropriateness	AI generates inappropriate content	4
Socio-emotional Development	Reduced ability to form real relationships	3

3.2 Evolution of Concerns: From Beginning to End of Semester

In the Week 18 questionnaires and interviews, the researcher asked parents whether their initial concerns had changed.

P7 (S7's mother) - Change in concerns

P7's greatest concern at the beginning of the semester was data privacy. In the Week 18 interview, she said:

> "At the beginning of the semester, I was worried about my child's conversations being recorded. But during the semester, I observed and also asked my child a few times. He said the teacher reminds them to close the window every time. I also saw that you didn't ask children to use their real names. So now I'm not so worried anymore. I trust that you are paying attention."
> -P7, Week 18 interview

P7's concern shifted from "highly worried" to "relatively reassured." The key factor in this shift was that she observed the teacher's specific actions (reminders to close windows, not using real names) and communicated with her child.

P2 (S2's mother) - Change in concerns

P2's greatest concern at the beginning of the semester was screen time. In the Week 18 questionnaire, she wrote:

> "At the beginning of the semester, I was worried about too much screen time. But later I learned that Al lessons are only once a week, and AI is not used throughout the entire lesson. My child also mentioned Chinese class at home, saying 'Today we used Al to draw a menu' it sounded quite fun. I think occasional use is acceptable."
> -P2, Week 18 questionnaire

P2's concern shifted from "strong opposition" to "conditional acceptance." The key factors were that she learned the details of the curriculum design (frequency and duration of AI use) and heard positive feedback from her child.

A broader look at screen time concerns across all parents: At Week 1, six parents expressed concern about screen time (see Table 6.1). By Week 18, among the nine parents who returned the questionnaire, three remained somewhat concerned, four were no longer concerned, and two did not return the questionnaire (so their current stance is unknown). What explained the shift among the four parents who became less concerned? Their open-ended responses pointed to two factors. First, they learned that AI use was limited to approximately 20 minutes per week (one 50-minute lesson with only partial AI use). Second, they observed that their children did not seem "glued to the screen" but rather engaged in active speaking and gesturing during AI tasks. One parent wrote: "I thought Al lessons would be like staring at a tablet. But my child showed me he uses hand gestures, turns to his partner, and talks out loud. It's not passive screen time." This distinction between passive consumption (e.g., watching videos) and active production (e.g., speaking to AI) appears to be important for parents' acceptance. Future parent communication might emphasize this distinction explicitly.

P11 (S11's mother) - Change in concerns

P11's greatest concern at the beginning of the semester was socio-emotional development. In the Week 18 interview, she said:

> "I'm still a little worried. But I observed my child-the way she speaks to the Al is different from the way she speaks to people. She knows the AI is not human. She doesn't treat the AI as a friend. So my worry might be a bit unnecessary. She is smarter than I thought."
> -P11, Week 18 interview

P11's concern shifted from "worry" to "partial reassurance." The key factor was that she observed her child's "correct" way of using the AI-S11 consistently perceived the AI as a tool, not as a substitute for real relationships.

P3 (S3's mother) - Change in concerns

P3's greatest concern at the beginning of the semester was technology dependence. In the Week 18 interview, she said:

> "I observed that my child occasionally uses the AI to look up words at home. But she doesn't depend on it-she thinks for herself first, and only uses the AI if she can't figure it

> out. I think this is good. She knows the Al is a tool, not an answer machine."
> -P3, Week 18 interview

P3's concern shifted from "worrying about dependence" to "observing responsible use." The key factor was that she saw her child's autonomy when using the AI ("thinking for herself first").

Table 6.2 Parent attitude change from Week 2 to Week 18

Parent	Primary Concern (Week 2)	Attitude at Week 2	Attitude at Week 18	Key Factor in Change
P3 (S3's mother)	Technology dependence	Cautiously supportive	Supportive	Observed child thinking independently
P7 (S7's mother)	Data privacy	Highly worried	Relatively reassured	Saw teacher's protective actions; child's confidence increased
P11 (S11's mother)	Socio-emotional development	Worried	Partially reassured	Observed child knows AI is not human
P2 (S2's mother)	Screen time	Opposed (initially)	Conditionally accepting	Learned curriculum details; child positive feedback
P10	General AI opposition	Strongly opposed	Unchanged (opposed)	Fundamental philosophical opposition to Al in schools

Note: P3, P7, and P11 participated in interviews at both Week 2 and Week 18. P2 and P10 completed only questionnaires. The table includes only parents whose attitudes changed (or notably did not change) from beginning to end of semester.

3.3 Unresolved Concerns

Not all parents' concerns were resolved. In the Week 18 questionnaire, one parent (P10) still expressed persistent worry:

> "I still think AI should not be in the classroom. It's not about this particular class; I oppose any AI in schools. Children should learn from real people, not machines. I don't see any benefit of AI. But I didn't want my child to be singled out, so I didn't withdraw this semester. If there are Al lessons next semester, I might consider switching classes."
> -P10, Week 18 questionnaire

P10's response reminds the researcher that not all parents can be persuaded. No matter how many protective measures teachers take, no matter how carefully the

curriculum is designed, some parents will fundamentally oppose Al in education. This is a reality teachers need to accept when promoting Al teaching. Having detailed parents' concerns at the beginning of the semester, I now examine how these concerns changed and what acceptance looked like by Week 18.

4 Theme 3: "I See Changes" - Parents' Acceptance and Recognition

4.1 Changes Parents Observed in Their Children

In the Week 18 questionnaires and interviews, the researcher asked parents: "Over this semester, have you observed your child mentioning the AI or Chinese lessons at home? Have there been any changes?" Multiple parents reported positive changes.

P7 (S7's mother) - Observation

> "He used to be unwilling to speak Chinese. At home, when we spoke Chinese to him, he often answered in English. But this semester, he sometimes voluntarily speaks Chinese for example, at dinner, saying 'Mom, this is very delicious.' He also explained to me how to say 'traffic light,' saying 'The AI didn't understand, I taught it.' He was very proud when teaching me. I think this is very good; he has gained confidence."
> -P7, Week 18 interview

P7's observation of "gained confidence" is consistent with the change observed in classroom observations of S7 from "not daring to speak" to "loudly correcting the AI." P7's narrative provides external validation for this study's classroom observations.

P3 (S3's mother) - Observation

> "Sometimes she tells me, 'Today the AI didn't understand again, and I taught it a word.' She describes in detail how she taught it what words she used, what gestures. I think in the process of 'teaching' the AI, she is also reviewing. She seems to have realized this herself, because she says, 'The process of teaching it is also learning'-those were her exact words."
> -P3, Week 18 interview

The phrase "the process of teaching it is also learning" mentioned by P3 is exactly what S3 said in the Week 18 interview. P3's observation confirms that S3 indeed brought this understanding home and shared it with her parents.

P11 (S11's mother) - Observation

> "She wasn't particularly excited about the AI. She just said very calmly, 'Today we used AI to make a menu,' 'The AI got it wrong, so I fixed it.' I think this is very good-she treats the AI as a normal tool and doesn't think it's anything special. This shows she is not 'fooled' by technology; she knows how to use it."
> -P11, Week 18 interview

P11's observation of "calm use" is consistent with S11's behavioral pattern in classroom observations of consistently perceiving the AI as a tool. Pll's evaluation of this pattern is positive ("not 'fooled' by technology").

4.2 Parents' Reassessment of Al's Value

In the Week 18 questionnaire, the researcher asked: "What is your current attitude toward AI-assisted Chinese lessons?" (open-ended question). Below are some representative responses:

> "From worry to support. Because I saw changes in my child-he is more willing to speak Chinese. If it's just one more tool, why not?"
> -P7, Week 18 questionnaire

> "I still think it shouldn't be too much. Once a week is acceptable. If it became two or three times a week, I would oppose it."
> -P2, Week 18 questionnaire

> "I think Al is helpful but not essential. A good teacher is more important than AI. AI can be a supplement."
> -P9, Week 18 questionnaire

> "I used to worry about technology dependence; now I'm less worried. Because my child knows the Al is not omnipotent; it makes mistakes. This has actually taught my child 'not to trust AI completely'-I think this is a very important skill."
> -P3, Week 18 questionnaire

P3's "taught my child 'not to trust AI completely"' is a highly insightful observation. She perceives the Al's "imperfection" as an educational opportunity students learn to evaluate Al outputs critically rather than accepting them blindly. This is a core component of "digital citizenship education."

4.3 Parents' Suggestions for Future AI Use

In the Week 18 questionnaire, the researcher asked: "If you could give one piece of advice to the school/teacher about AI use, what would it be?" Below are some representative suggestions:

> "Continue to be transparent. Tell us what Al is used, how it is used, and how data is handled. We are not opposed to technology, but we need to know."
> -P4, Week 18 questionnaire

> "Don't use it too much. Al is a tool; it cannot replace teachers. Children need to speak with real people; they need teachers' emotional support."
> -P6, Week 18 questionnaire

> "Teach children how to use Al safely. Not just in Chinese class, but in all classes. This is a skill they will need in the future."
> -P11, Week 18 questionnaire

> "If possible, give parents an 'observation day' let us see how AI lessons are actually conducted. Reading a letter is not enough; we want to see it with our own eyes."
> -P1, Week 18 questionnaire

Pl's "observation day" suggestion is a good idea. The researcher wrote in the Week 18 reflective journal:

> "Parents said they want to see how AI lessons are taught. I think this is a good suggestion. Next semester, I could show a short classroom video clip (without showing students' faces) at a parent meeting, or invite a few parents to observe. That would be more effective than writing ten letters."
> -Teacher reflective journal, Week 18

5 Chapter Summary

This chapter has presented three main findings of this study regarding "privacy, ethics, and digital citizenship education."

First, teachers need to adopt multi-layered, sustained ethical practices in AI classrooms. These practices include: parental informed consent at the beginning of the semester (detailed parent letter + phone conversations), privacy protection measures in the classroom (no real names, clearing dialogues after class, data encryption), timely handling of unexpected ethical dilemmas (e.g., students typing personal information, AI generating inappropriate content), and sustained reflexive reflection (e.g., recording one's own emotional responses and potential biases). This study summarizes a "AI Classroom Ethical Practices: Teacher's Do Not Do List" that can serve as a reference for other teachers.

Second, parents' concerns about introducing AI into the Chinese classroom are multidimensional, including data privacy, screen time, technology dependence, content appropriateness, and socio-emotional development. Among these, data privacy was the most common and strongest concern. These concerns evolved over the semester-some parents' concerns shifted from "highly worried" to "relatively reassured" or "conditionally accepting." The key factors in this shift were: parents observed the teacher's specific protective actions, learned details of the curriculum design, heard positive feedback from their children, and saw their children using Al responsibly. However, not all parents' concerns were resolved some parents fundamentally opposed AI in education.

Third, parents observed positive changes in their children, including greater willingness to speak Chinese, increased confidence, and the development of critical AI use skills ("not trusting AI completely"). Some parents consequently reassessed the value of AI-from "worry" to "support" or "conditional acceptance." Parents' suggestions for future AI use included: maintaining transparency, controlling frequency of use, teaching children safe AI use skills, and providing parents with "observation day" opportunities.

Together, these findings reveal that introducing AI into the classroom is not merely a technical issue but also an ethical issue, a communication issue, and a trust issue. Teachers' ethical practices cannot stop at "writing a letter" or "signing a consent form"-they require sustained action and reflection throughout the semester. Parents' concerns should not be treated as "obstacles" or "nuisances" but as valuable

feedback for improving instructional design and communication strategies. When parents see positive changes in their children and observe teachers' responsible actions, their attitudes can shift from skepticism to support-but this shift takes time and evidence.

Chapter 7
Discussion and Conclusion

Huang Xiao

Abstract: This chapter is the final chapter of the thesis, aiming to engage in theoretical dialogue between the research findings presented in the previous three chapters and the literature review in Chapter 2, answering the question "What do the findings of this study mean?"

1 Summary of Research Findings

1.1 Answer to RQ1: Human-Machine Negotiation of Meaning

RQ1: In an AI-assisted junior high CSL classroom, what kinds of meaning negotiation behaviors occur between students and the AI? How does the AI's "intentional error-making" setting affect students' language output and modification? This study found:

First, students' initial responses to the AI's "failure to understand" followed an evolutionary trajectory from "surprise" to "expectation" (tension-silence stage → excitement-correction stage → calm-strategic stage). This evolution indicates that the ability to negotiate meaning develops gradually through interactional experience.

Second, students used five meaning negotiation strategies: repetition, paraphrase, simplification, code-switching to English, and multimodal assistance. Strategy choice was related to language proficiency-high-proficiency students preferred paraphrase

Huang Xiao • Krirk University, Bangkok, Thailand
e-mail: huang.xiao@krirk.ac.th

© The Author(s). Published in the USA.
https://doi.org/10.63408/979-8-9928364-4-8

and simplification, while weaker-proficiency students preferred repetition and code-switching.

Third, the AI's "intentional error-making" setting successfully created abundant opportunities for meaning negotiation, transforming students from "the corrected" to "the correctors." Students experienced a sense of competence in the process of "teaching the AI" and developed metacognitive understanding "the process of teaching it is also learning."

1.2 Answer to RQ2: AI Role Perception

RQ2: How do students perceive the AI's role in their Chinese learning? How does this role perception evolve over time and across interactional situations? This study found:

First, students' perceptions of Al roles were diverse, including six types: tool, peer in need, fallible opponent, coach, teacher, and others (e.g., "mirror"). Role perception was related to language proficiency-high-proficiency students tended toward tool perception, while weaker-proficiency students tended toward opponent or peer perception.

Second, Al roles are fluid. The same student, in the same lesson, could assign different roles to the AI depending on changes in the interactional context. Key events triggering role switching included the Al's correct response, the Al's first mistake, the Al's repeated mistakes, and the student getting stuck.

Third, Al role perception evolved over the semester. Students with lower language proficiency showed more dramatic evolution in role perception. S7's evolution trajectory was: testing machine → fallible opponent → game partner → learning partner; S3's evolution trajectory was: peer in need → stable peer; S11's evolution trajectory was: tool → more efficient tool.

1.3 Answer to RQ3: Privacy, Ethics, and Digital Citizenship Education

RQ3: What privacy and ethical protection strategies does the teacher employ in AI-assisted teaching? What narratives of concern and acceptance do parents exhibit regarding the introduction of AI into the Chinese classroom? This study found:

First, teachers need to adopt multi-layered, sustained ethical practices, including: parental informed consent at the beginning of the semester (detailed parent letter + phone conversations), privacy protection measures in the classroom (no real names, clearing dialogues after class, data encryption), timely handling of unexpected ethical dilemmas, and sustained reflexive reflection.

Second, parents' concerns were multidimensional, including data privacy, screen time, technology dependence, content appropriateness, and socio-emotional devel-

opment. Data privacy was the most common and strongest concern. These concerns evolved over the semester some parents' concerns shifted from "highly worried" to "relatively reassured" or "conditionally accepting." The key factors in this shift were: parents observed the teacher's specific protective actions, learned details of the curriculum design, heard positive feedback from their children, and saw their children using AI responsibly.

Third, parents observed positive changes in their children, including greater willingness to speak Chinese, increased confidence, and the development of critical AI use skills. Some parents consequently reassessed the value of AI.

2 Dialogue with Theory

2.1 Extending the Interaction Hypothesis and Output Hypothesis

How do the findings of this study dialogue with the Interaction Hypothesis? Long's (1996) Interaction Hypothesis, based on empirical research on human-human interaction, emphasizes that negotiation of meaning promotes acquisition. Data show that human-machine meaning negotiation shares similarities with human-human meaning negotiation but also has differences.

Similarities: Meaning negotiation did occur in student-AI interactions (clarification requests, confirmation checks, repetition, paraphrase, etc.), and these negotiations triggered students' output modification. This is consistent with the mechanism Long described.

Differences: The AI's "predictable errors" triggered students' output modification more effectively than the "natural errors" in human-human conversation. In this study, the Al's "intentional errors" were predictable-students quickly learned to anticipate where the AI would make mistakes. This predictability lowered students' anxiety, making them more willing to correct proactively. In human conversation, the other person's errors are unpredictable, and learners cannot "prepare in advance" correction strategies. This finding suggests that predictability may be a unique advantage of human-machine meaning negotiation. The Al's "silliness" is not a technical flaw but a pedagogical design-it creates a "safe" environment for meaning negotiation, where students know the Al will make mistakes, know what types of mistakes, and can therefore engage in negotiation relaxedly and proactively.

How do the findings of this study dialogue with the Output Hypothesis? Swain's (2005) Output Hypothesis proposed three functions: noticing trigger, hypothesis testing, and metalinguistic reflection. The findings of this study provide empirical evidence for the operation of these three functions in human-machine interaction:

Noticing trigger: The Al's "failure to understand" triggered students' noticing. S3 said in an interview, "Did I not say 'right' clearly?" this is a manifestation of noticing being triggered.

Hypothesis testing: Students used the Al's responses to test their language hypotheses. S11 said, "The AI understood, so it should be correct" she used the Al's understanding as evidence of her own linguistic correctness.

Metalinguistic reflection: Students' deepest metalinguistic reflection occurred in the process of "teaching the AI." S7 said, "The process of teaching it is also learning"-he realized that the process of explaining language rules to the AI was itself the process of consolidating and understanding those rules.

The unique contribution of this study is revealing that "teaching the AI" as a special form of output may promote metalinguistic reflection more strongly than "outputting to the AI." When students merely "output," they focus on "can the Al understand?" When students "teach" the AI, they focus on "can I explain clearly?" this requires a higher level of metalinguistic processing.

2.2 Re-examining the Affective Filter Hypothesis

Krashen's (1982) Affective Filter Hypothesis argues that low anxiety promotes acquisition. The findings of this study support this hypothesis but offer an important supplement: the AI's "non-judgmental" quality does lower anxiety, but the effect of anxiety reduction varies by individual and over time.

For S7 (weaker proficiency), the Al's low-anxiety environment had a significant effect he moved from "not daring to speak" to "loudly correcting the AI." For S11 (high proficiency), the Al's low-anxiety environment had little effect on her she was not very anxious to begin with. For S3 (intermediate-high proficiency), the Al's low-anxiety environment made her feel "not nervous but finds it troublesome"-anxiety was lowered but did not translate into positive affect.

More importantly, observations indicate that anxiety reduction is not a one-time achievement. At the beginning of the semester, students were nervous about the AI (because it was unfamiliar); in the middle of the semester, they relaxed; but when the AI's "degree of silliness" exceeded their expectations (e.g., the unpredictable errors in Week 9), anxiety returned. This suggests that teachers need to continuously attend to students' affective states and adjust the Al's "degree of silliness" based on students' reactions.

2.3 Extending the AI Role Framework

Luckin et al. (2016) proposed four roles for Al in education (tutor, coach, peer, tool), but this framework tends to treat Al roles as fixed. The findings of this study propose two important extensions to this framework.

Extension 1: Role fluidity

The findings reveal that the same student, in the same lesson, could assign different roles to the AI depending on changes in the interactional context. S7 experienced

switches from "tool" to "opponent" to "game partner" within 15 minutes in Week 5. This indicates that Al roles are not technologically predetermined but are interactional achievements dynamically constructed by students in moment-to-moment interaction. The implication of this finding for Al instructional design is that teachers should not expect students' perceptions of AI roles to be "stable" or "presettable." Instead of asking "What role should this Al be designed to play?", it is more useful to ask "Under what conditions will students perceive the AI as what role?" The latter is a more dynamic, context-sensitive design question.

Extension 2: "Fallible peer" as a distinct role type

Tsai (2021) proposed the concept of "AI as a mistake-prone conversation partner." The findings of this study provide rich qualitative evidence for this concept and further distinguish two subtypes of "fallible peer":

- **Peer in need (S3's perception):** The AI is an object with limited abilities that needs to be taught. The student's affective experience is patience and satisfaction.
- **Fallible opponent (S7's perception):** The AI is an object that can be "defeated." The student's affective experience is excitement and pleasure.

Although both subtypes are based on the AI's "mistakes," the student's stance differs-one is a "helper," the other is a "competitor." This distinction has implications for instructional design: if the goal is to cultivate patience and explanation skills, the "peer in need" setting may be more appropriate; if the goal is to stimulate participation motivation and sense of competence, the "fallible opponent" setting may be more effective.

2.4 Applying the RRI Framework to Educational AI

Von Schomberg's (2013) RRI framework emphasizes five dimensions: anticipation, reflexivity, inclusion, responsiveness, and sustainability. The findings of this study provide empirical cases for the application of this framework to educational AI.

Inclusion dimension: This study involved parents in the decision-making process regarding AI use through parent letters, phone conversations, and informed consent. P7 said, "I trust that you are paying attention" this suggests that when parents are "included," their concerns are more easily alleviated.

Responsiveness dimension: The researcher adjusted teaching practices in response to parent feedback (e.g., P2's concern about screen time) and unexpected classroom events (e.g., S8 typing his real name, AI generating inappropriate content), such as adding reminders and enabling stricter content filtering. This indicates that RRI's "responsiveness" is not a one-time event but a continuous process.

Reflexivity dimension: The researcher documented the process of ethical decision-making and emotional responses through reflective journals. This reflexivity serves not only as researcher self-monitoring but also as a mechanism for ensuring research quality.

The findings of this study suggest that the operationalization of the RRI framework in educational AI needs to be more concrete. Teachers need not abstract "responsible innovation" principles but actionable "how-to" guides-for example, how to handle a student inadvertently typing personal information? How to respond to AI generating inappropriate content? The "Teacher's Do Not Do List" summarized in this study is an attempt in this direction.

One theoretical perspective notably absent from the original literature review but highly relevant to this study's core finding - "teaching the AI is also learning" - is Vygotskian sociocultural theory. Vygotsky's (1978) concept of the Zone of Proximal Development (ZPD) describes the space between what a learner can do independently and what they can do with assistance. In traditional ZPD conceptualizations, the "more knowledgeable other" is typically a teacher or peer. However, in this study's "teach the AI" tasks, the roles are reversed: the student becomes the "more knowledgeable other" and the AI becomes the learner. This reversal - which Lantolf and Poehner (2014) might call a "reverse ZPD" deserves further theoretical attention. The finding that students develop metalinguistic awareness through explaining language rules to the Al resonates strongly with sociocultural arguments about the internalization of knowledge through teaching others. Future research might explicitly examine Al-mediated peer tutoring through a sociocultural lens.

3 Theoretical Contributions

Based on the foregoing dialogue with theory, this study proposes the following theoretical contributions:

3.1 Conceptualizing "Human-Machine Negotiation of Meaning"

This study proposes "human-machine negotiation of meaning" as an extended concept of the Interaction Hypothesis in the AI era. Compared to "human-human negotiation of meaning," human-machine negotiation of meaning has the following characteristics:

Table 7.1 Comparison of Meaning Negotiation

This conceptualization can help future research more precisely describe and compare meaning negotiation in human-human versus human-machine interaction.

Dimension	Human-Human Negotiation	Human-Machine Negotiation
Predictability of errors	Low (other's errors unpredictable)	High (Al's "intentional errors" can be designed)
Affective risk	High (social anxiety, face concerns)	Low (AI does not judge, does not mock)
Role relationship	Usually asymmetrical (NS-NNS)	Designable (student can be the "corrector")
Immediacy of feedback	Immediate	Immediate
Consistency of feedback	Inconsistent (varies by individual)	Consistent (AI responds according to rules)

3.2 Proposing the Concept of "Role Fluidity"

This study proposes "role fluidity" as a supplement to Luckin et al.'s (2016) Al role framework. Role fluidity refers to the phenomenon in which students dynamically assign different roles to the AI depending on changes in the interactional context during the process of interacting with the AI. The core claim of role fluidity is that the Al's role is neither technologically predetermined nor fixedly assigned by students, but is co-constructed by the interactional behaviors of both parties (student and AI) in moment-to-moment interaction. This concept shifts the analytical focus from "what role is the AI?" to "under what conditions do students perceive the AI as what role?", providing a more dynamic perspective for AI instructional design.

3.3 The "From Opponent to Partner" Evolution Trajectory Model

Based on S7's evolution trajectory (testing machine → fallible opponent → game partner → learning partner), this study proposes a hypothetical evolution trajectory model for future research to test:

Novice Stage → Confrontation Stage → Game Stage → Collaboration Stage

- **Novice Stage:** Students do not understand the Al's capabilities and limitations, feeling tense and uncertain.
- **Confrontation Stage:** Students discover the Al makes mistakes and begin correcting it, perceiving the Al as an "opponent."
- **Game Stage:** Students enjoy the process of correcting the AI, treating the interaction as a game and sharing the Al's "silliness" with peers.
- **Collaboration Stage:** Students perceive the AI as a "learning partner" they can cooperate with, using the AI calmly and strategically.

This model is not "a necessary stage that all students will experience" but "a typical trajectory that may occur under certain conditions." Future research can test

the applicability of this model across different age groups, language proficiency levels, and AI settings.

3.4 An Operationalization Framework for "Teacher Ethical Practices"

Based on the experience of this study, a preliminary "Operationalization Framework for AI Classroom Ethical Practices" is proposed, comprising four dimensions:

Table 7.2 Operationalization Framework

Dimension	Operationalization Suggestions
Preparation	Write a detailed parent letter (including "Frequently Asked Questions"); obtain informed consent; prepare backup activities (in case of technical failures)
Classroom implementation	Remind students not to use real names; circulate to monitor AI output; handle unexpected events promptly; clear dialogues after class
Ongoing communication	Provide regular updates to parents on AI use; collect parent feedback; adjust practices based on feedback
Reflexive reflection	Document ethical decisions and emotional responses; actively seek disconfirming evidence; invite peer review

This framework is preliminary and context-specific; future research can test and revise it in different contexts.

4 Pedagogical Recommendations

Based on the findings of this study, the following recommendations are offered for international school CSL teachers who wish to introduce AI into second language classrooms:

4.1 Regarding AI "Personality" Design

Recommendation 1: Consider using a "fallible AI" rather than a "perfect AI." Analysis shows that the AI's "intentional errors" created opportunities for meaning ne-

gotiation, transforming students from "the corrected" to "the correctors." If the AI is always correct, students may simply receive passively without actively thinking.

Recommendation 2: The Al's "degree of silliness" needs careful design. Too smart → fewer negotiation opportunities; too silly → student frustration. This study's experience suggests that AI errors should be predictable (e.g., always reversing directions) rather than random and unpredictable. Predictable errors allow students to "learn" the AI's error patterns and proactively prepare correction strategies.

Recommendation 3: Adjust the Al's "degree of silliness" according to students' language proficiency. For weaker-proficiency students, Al errors should be simpler and more predictable; for high-proficiency students, the AI can be "smarter," or students can be allowed to set the Al's role themselves.

4.2 Regarding Activity Design

Recommendation 4: AI use should be "short, distributed, and supervised." This study used one AI lesson per week, with AI use lasting approximately 20-25 minutes per AI lesson. Do not give entire lessons to the AI-students need real-person interaction, writing practice, and cultural experiences.

Recommendation 5: Design "teach the AI" tasks. Rather than having students "output to the AI," have them "teach" the AI. For example, have the AI play the role of a "foreigner who just started learning Chinese," and students need to explain a word or grammar rule. The process of "teaching" promotes metalinguistic reflection more than the process of "speaking."

Recommendation 6: Prepare no-Al backup activities (e.g., pair-based role-play using printed dialogue cards or vocabulary bingo). Technical failures are inevitable. When the AI lags, the network disconnects, or the AI generates inappropriate content, teachers need to be able to switch immediately to a backup activity without disrupting the lesson.

4.3 Regarding Privacy and Ethics

Recommendation 7: Write a detailed parent letter and maintain ongoing communication. The parent letter should include: why Al is used, what Al is used, how data is handled, and parents' rights. Do not write it only once-provide regular updates during the semester and collect parent feedback.

Recommendation 8: Remind students "not to type real names" at the beginning of each Al lesson. One reminder at the beginning of the semester is insufficient. This study's experience suggests that a 10-second reminder before each Al activity can effectively reduce incidents.

Recommendation 9: Circulate to monitor Al output. Do not assume that AI content filtering is perfect. Teachers need to circulate while students use AI to promptly detect and address inappropriate content.

Recommendation 10: Clear dialogue logs after class. Ensure student conversations are not saved. This not only protects privacy but also lets students know that "the AI won't remember what I said wrong"-which further lowers anxiety.

4.4 Regarding Emotional Support

Recommendation 11: Attend to students' affective responses to the AI. Some students (e.g., S7) enjoy interacting with the AI, some (e.g., S3) find it "troublesome," and some (e.g., S11) remain neutral. If a student shows persistent frustration or anxiety, the teacher needs to intervene possibly by adjusting the Al's settings or providing additional support.

Recommendation 12: Do not assume that "all students like the AI." Some students may prefer speaking with real people. Provide choices-for example, allow students to choose whether to complete a task by conversing with the AI or with a peer.

4.5 BOX: Quick Reference - 5 Things Teachers Can Do Tomorrow

- **Make the Al intentionally "silly" (but predictably so):** Design AI to make predictable errors (e.g., always reversing directions) rather than random ones. This creates safe, low-anxiety opportunities for students to become "correctors."
- **Remind students "Don't type your name" before every Al lesson:** One reminder at the beginning of the semester is not enough. A 10-second reminder before each Al activity prevents privacy incidents.
- **Design "teach the AI" tasks, not just "talk to AI" tasks:** Have AI play a "foreigner who just started learning Chinese." Students must explain words or grammar rules. The process of teaching promotes deeper learning than simply speaking.
- **Prepare a no-Al backup activity:** Technical failures happen. Keep a simple backup ready (e.g., pair-based role-play using printed dialogue cards) so Al lessons never collapse into chaos.
- **Send a parent letter with FAQs - and keep communicating:** Including why AI, which AI, how data is handled, and parents' rights. Then provide updates during the semester and collect feedback. Transparency builds trust.

5 Research Limitations

This study has the following limitations, which readers should consider when interpreting the findings:

5.1 Methodological Limitations

Limitation 1: Influence of the researcher's dual role. The researcher of this study was simultaneously the class's Chinese teacher and curriculum designer. Although this identity brought advantages in terms of deep field access, it may also have introduced bias-students may have given the answers "the teacher wanted to hear" in interviews; the researcher may have had emotional investment in their own curriculum design, affecting the objectivity of observations and interpretations. This study addressed this limitation through strategies such as reflexivity journals, peer review, and actively seeking disconfirming evidence, but the influence cannot be completely eliminated.

Limitation 2: Small sample size and context-specificity. This study involved only one class (14 students) in the specific context of a junior high CSL classroom at an international school. The specificity of this context (multicultural backgrounds, English as the main language of instruction, HSK Level 3 proficiency) means that the findings cannot be directly generalized to other contexts (e.g., public schools, adult learners, other languages). The goal of this study is not "statistical generalization" but "theoretical generalization" and "contextual resonance"-readers need to judge whether "this study's context is similar to my context."

Limitation 3: Limited time span. Eighteen weeks is one semester, sufficient to observe students' trajectories of change but insufficient to observe longer-term effects (e.g., whether students still remember the experience of "teaching the AI" one year later). Longer-term effects require future longitudinal research.

5.2 Technical Limitations

Limitation 4: Limitations of the Al tool. This study used the voice version of ChatGPT, a general-purpose AI not specifically designed for language teaching. Al tools specifically designed for language teaching may have different functions and limitations. Furthermore, Al technology is iterating rapidly-the version of ChatGPT used in this study may soon be replaced by newer versions. The findings need to be re-examined as technology evolves.

Limitation 5: Impact of technical failures. During the research process, there were several instances of network latency, slow Al responses, or the AI "freezing." These technical failures affected the flow of instruction and may also have affected

students' experiences. In environments with more stable technology, students' experiences might differ.

5.3 Theoretical Limitations

Limitation 6: Specificity of the theoretical framework. This study primarily used the Interaction Hypothesis, Output Hypothesis, Affective Filter Hypothesis, Al role framework, and RRI framework as theoretical tools. Other theoretical perspectives (e.g., sociocultural theory, activity theory, actor-network theory) might reveal different aspects of the findings. This study is not a "comprehensive" theoretical analysis but a "focused" theoretical dialogue.

Limitation 7: Concepts are still in development. The concepts proposed in this study-"human-machine negotiation of meaning," "role fluidity," "from opponent to partner evolution trajectory"-are preliminary and exploratory. They need to be tested, revised, or falsified in different contexts through future research.

6 Future Research Directions

Based on the findings and limitations of this study, the following future research directions are proposed:

6.1 Comparative Research

Direction 1: Comparative study of human-machine vs. human-human meaning negotiation. This study identified some characteristics of human-machine meaning negotiation (e.g., predictability of errors, low affective risk), but do these characteristics actually promote acquisition? Future research could compare two groups of learners (one negotiating with AI, one negotiating with a human) in an experimental environment and measure their language progress.

Direction 2: Comparative study of different AI "personalities." This study used a "fallible AI" setting. If a "perfect AI" (never making mistakes) were used, how would students' meaning negotiation behaviors and role perceptions differ? Future research could design comparative experiments to compare the effects of "silly AI" vs. "smart AI" on learning outcomes.

Direction 3: Comparative study of different age groups. This study focused on junior high students (ages 12-14). Do elementary students, high school students, and adult learners perceive AI roles differently? Do their meaning negotiation strategies differ? Future research could conduct similar classroom observations across different age groups.

6.2 Longitudinal Research

Direction 4: Longitudinal tracking of long-term effects. Students reported positive changes after 18 weeks, but are these changes durable? One semester later, one year later, do students still remember the experience of "teaching the AI"? Does their Al literacy continue to develop? Future research could conduct longer-term tracking.

Direction 5: Testing the evolution trajectory model. This study proposed a "from opponent to partner" evolution trajectory model based on S7's case. Does this model apply to more students? Are there other types of evolution trajectories (e.g., "from tool to partner" or "from peer to tool")? Future research could test and revise this model with larger samples.

6.3 Design Research

Direction 6: Design principles for AI "degree of silliness." How silly should the AI be? What kinds of errors are "pedagogically valuable errors"? What kinds of errors become "frustrating errors"? Future research could systematically manipulate the types and frequencies of AI errors, observe students' responses, and distill design principles.

Direction 7: Research on teacher professional development. What new knowledge and skills do teachers need in AI classrooms? How can teachers be trained to design Al activities, handle ethical dilemmas, and support students' emotional needs? Future research could develop and evaluate teacher training programs.

6.4 Ethics and Policy Research

Direction 8: Cross-cultural comparison of parental concerns. The parents in this study came from diverse cultural backgrounds (Western and Asian). Do parents from different cultural backgrounds have different attitudes toward introducing AI into classrooms? Future research could conduct cross-cultural comparisons to inform policy-making in international schools.

Direction 9: Development and evaluation of Al ethics curricula. This study suggests "teaching students how to use AI safely." Future research could develop and evaluate "AI literacy" or "digital citizenship" curricula for junior high students.

7 Conclusion

This study is a qualitative investigation of AI-assisted junior high CSL teaching. Through 18 weeks of classroom observations, student interviews, parent questionnaires, teacher reflective journals, and other multi-source data, it explored three core questions: meaning negotiation behaviors between students and the AI, students' perceptions of Al roles and their evolution, and teacher ethical practices and parental concerns.

This study found that the Al's "intentional error-making" setting successfully created abundant opportunities for meaning negotiation, transforming students from "the corrected" to "the correctors," and developing their language abilities and communicative strategies in the process of "teaching the AI." Students' perceptions of Al roles were diverse, fluid, and evolved over time and across situations students with lower language proficiency showed more dramatic evolution in role perception. Teachers' ethical practices need to span the entire semester, including preparation, classroom implementation, ongoing communication, and reflexive reflection. Parents' concerns were multidimensional, but through transparent communication and observing positive changes in their children, some parents' worries were transformed into support.

The central argument of this study is that the value of AI in the language classroom lies not in its "perfection" but precisely in its "imperfection." An AI that never makes mistakes is a perfect tool, but not a good teacher. A good teacher-whether human or AI-should create opportunities for students to think, try, make mistakes, and correct. The Al's "silliness" is not a technical flaw but a pedagogical design-it creates opportunities for "role reversal," transforming students from passive recipients into active "correctors" and "teachers."

Of course, Al is not a panacea. It cannot replace the teacher's emotional support, cultural explanation, and deep understanding of individual students. The optimal role of AI is "assistant" a language partner that can accompany students 24/7, never loses patience, but also makes mistakes. True learning occurs when students "teach" the AI-when they strive to explain, revise repeatedly, and try every means to make the AI understand, they are also teaching themselves.

This study is one of the early explorations of AI entering the language classroom. In today's rapidly iterating technological landscape, specific tools may become obsolete, and specific activity designs may need updating. But the core findings of this study the value of meaning negotiation, the existence of role fluidity, the importance of ethical practices, and the mechanism of "teaching the AI" as learning-may have more lasting significance. It is hoped that these findings can serve as a reference for other teachers, researchers, and policymakers, collectively exploring a responsible path for the integration of AI and education.

A final word for language teachers who are not researchers. If you take away only one thing from this book, let it be this: a perfect AI one that never makes mistakes is not your students' best learning partner. An AI that makes predictable mistakes, an Al that sometimes "doesn't understand," an AI that needs to be taught this imperfect Al creates something precious: a moment when your student becomes

the expert, the corrector, the teacher. In that moment, anxiety drops, confidence rises, and learning happens. So do not be afraid to let your AI be a little silly. Design its errors carefully (predictable, not random). Protect your students' privacy (no real names, clear the logs). Talk to parents honestly and often. And then step back and watch what happens. You may be surprised, as I was, by which students flourish. The quiet one who never raises her hand may become the loudest corrector of the AI. That is the power of a silly language partner.

References

Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *Al and Ethics*, 2(3), 431-440.

Berger, R. (2015). Now I see it, now I don't: Researcher's position and reflexivity in qualitative research. *Qualitative Research*, 15(2), 219-234.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.

Braun, V., & Clarke, V. (2022). *Thematic analysis: A practical guide*. SAGE Publications.

Chen, X., Gao, Y., Guan, J., & Ryoo, J. (2024). AI application in foreign language literature: ChatGPT's impact and skill enhancement. *European Journal of Contemporary Education and E-Learning*, 2(2), 3-18.

Denzin, N. K., & Lincoln, Y. S. (Eds.). (2018). *The SAGE handbook of qualitative research* (5th ed.). SAGE Publications.

Duff, P. A. (2008). *Case study research in applied linguistics*. Lawrence Erlbaum Associates.

Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International Journal of Qualitative Methods*, 5(1), 80-92.

Gass, S. M., & Varonis, E. M. (1994). Input, interaction, and second language production. *Studies in Second Language Acquisition*, 16(3), 283-302.

Georgiou, G. P. (2026). Envisioning the futures of language education in the era of artificial intelligence. *Journal of Futures Studies*.

Godwin-Jones, R. (2022). Partnering with AI: Intelligent writing assistance and instructed language learning. *Language Learning & Technology*, 26(2), 5-24.

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2022). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education*, 32(3), 504-526.

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.

Kukulska-Hulme, A., & Shield, L. (2008). An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. *ReCALL*, 20(3), 271-289.

Lam, W. S. E. (2006). Culture and learning in the context of globalization: Research directions. *Review of Research in Education*, 30(1), 213-237.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE Publications.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). Academic Press.

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for Al in education*. Pearson.

OECD. (2022). *OECD privacy framework for Al in education*. OECD Publishing.

O'Connor, M. T. (2025). Als as fellow participants in the language game. *Al & Society*. https://doi.org/10.1007/s00146-025-02239-w

Pan, L., & Block, D. (2011). English as a "global language" in China: An investigation into learners' and teachers' language beliefs. *System*, 39(3), 391-402.

Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). SAGE Publications.

Penuel, W. R., Philip, T. M., Chang, M. A., Sumner, T., & D'Mello, S. K. (2025). Responsible innovation in designing AI for education: Attending to the "how," the "for what," the "for whom," and the "with whom" in a rapidly growing field. *Educational Researcher*. https://doi.org/10.3102/0013189X251234567

Pica, T. (1994). Research on negotiation: What does it reveal about second-language learning conditions, processes, and outcomes? *Language Learning*, 44(3), 493-527.

Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. Polity Press.

Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a framework for responsible innovation. *Research Policy*, 42(9), 1568-1580.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Newbury House.

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471-483). Lawrence Erlbaum Associates.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16(3), 371-391.

Tahmasbi, A., Esrafilian, M., Wright, J., Jeong, S., & Bera, A. (2026). Game Master LLM: Task-based role-playing for natural slang learning. *arXiv preprint, arXiv:2511.15504*.

Tsai, S.-C. (2021). Chatbot-assisted language learning: The effect of error types on EFL learners' engagement and negotiation of meaning. *Computer Assisted Language Learning*, 34(8), 1075-1099.

UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. UNESCO Publishing.
von Schomberg, R. (2013). A vision of responsible research and innovation. In R. Owen, J. Bessant, & M. Heintz (Eds.), *Responsible innovation* (pp. 51-74). John Wiley & Sons.
Warschauer, M., & Healey, D. (1998). Computers and language learning: An overview. *Language Teaching*, 31(2), 57-71.
Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE Publications.

Appendix

Huang Xiao

Appendix A: Letter to Parents

Letter to Parents: Making AI Your Child's "Smart Language Partner" in Chinese Learning

Dear Parent/Guardian,

Hello!

I am Xiao Huang, your child's Chinese teacher. In the upcoming semester, I will moderately introduce artificial intelligence (AI) assistance tools (such as ChatGPT voice version, Duolingo, Wordwall, etc.) into the Chinese as a second language classroom.

I understand that you may have some questions. Please allow me a few minutes to explain.

I. Why use Al in Chinese class?

Junior high school students are at a developmental stage where they crave interaction but fear making mistakes in front of peers. Al offers three irreplaceable benefits:

- **Reduces anxiety:** Al is an emotionless, non-judgmental practice partner. When you speak incorrectly to the AI, it only gives friendly prompts, without embarrassment.

Huang Xiao • Krirk University, Bangkok, Thailand
e-mail: huang.xiao@krirk.ac.th

© The Author(s). Published in the USA.
https://doi.org/10.63408/979-8-9928364-4-8

- **Increases practice:** Every student can converse with their own AI simultaneously, rather than having to wait their turn to practice with the teacher.
- **Provides immediate feedback:** Al can instantly point out pronunciation or grammar issues, helping students adjust in a timely manner.

AI will not replace the teacher. The teacher is responsible for designing the curriculum, explaining culture, grading assignments, and providing emotional support. The Al's role is like a "talking dictionary + 24/7 practice partner."

II. How do we protect your child's privacy?

We will strictly implement the following measures in the classroom:

1. **No collection of personal information:** Students do not need to log in when using AI, nor do they need to provide names, photos, home addresses, etc.
2. **No use of real names:** When conversing with the AI, students only use classroom codes or Chinese nicknames.
3. **No saving of conversation logs:** After each classroom activity, I will clear all conversation logs.
4. **No recording of faces:** If classroom video recordings are used for teaching research, faces will be blurred.
5. **Parental informed consent:** Without your written consent, your child will not participate in any Al activities.

III. Practical use in the classroom

A 50-minute Al-assisted lesson is typically arranged as follows:

- **First 5 minutes:** Al pronunciation challenge (whole class reads one word simultaneously)
- **Middle 15 minutes:** Students engage in role-play dialogues with the AI
- **Last 10 minutes:** AI-generated mini-games (e.g., whack-a-mole for character recognition)
- **Remaining time:** Traditional writing, group discussions, cultural activities

IV. What you can do

- **Ask questions anytime:** If you would like to know which app or privacy policy is being used on a specific day, I can provide the information at any time.
- **Opt out:** If you prefer that your child does not interact directly with the AI, I can arrange for your child to work with a peer instead of operating the device directly.
- **Optional home extension:** I can recommend some Al game websites that require no login and collect no data, for your child to voluntarily use after class.

V. My commitment

Your child's safety and growth are always the top priority. I will not sacrifice any privacy principle for the sake of pursuing "new technology." The sole purpose of introducing AI is to help those children who once found Chinese "too difficult and too boring" find the courage to speak and the joy of learning.

You are welcome to observe an AI lesson in the classroom and see the entire process for yourself.

Thank you for your trust and support!

Best regards,

Xiao Huang

Parent Reply Slip (please cut and return)

Informed Consent Form for Use of AI Tools in Chinese Class

I have read the above explanation and understand that AI tools will be used to assist Chinese language learning and that privacy protection measures are in place in the classroom.

☐ I consent to my child _____ (child's name) using AI tools for classroom activities under the teacher's supervision.
☐ I prefer that my child does not operate the Al directly but participates by collaborating with classmates.
☐ I would like to communicate further with the teacher. My contact information is: _____

Parent signature: _____
Date: _____

Appendix B: Parent Open-Ended Questionnaire (Week 1 and Week 18)

Week 1 Questionnaire

Parent Questionnaire on AI-Assisted Chinese Teaching (Beginning of Semester)

Dear Parent,

To better understand your thoughts, please take 5 minutes to answer the following questions. Your responses will help me improve my teaching. This questionnaire is anonymous; you do not need to provide your name.

1. Have you read the "Letter to Parents"?

☐ Yes
☐ No
☐ Partially

2. What are your main concerns about using AI in Chinese class? (Select all that apply)

☐ Data privacy (my child's conversations being recorded)
☐ Too much screen time
☐ My child becoming overly dependent on AI and not thinking for themselves
☐ AI may generate inappropriate content
☐ Reduced real-person interaction for my child
☐ Other: _____

3. What are your greatest hopes for using Al in Chinese class? (Select all that apply)

☐ My child becomes more willing to speak Chinese
☐ My child gets more practice opportunities
☐ My child becomes more interested in Chinese
☐ Other: _____

4. Is there anything else you would like to say to the teacher?

Thank you for your responses!

Week 18 Questionnaire

Parent Questionnaire on AI-Assisted Chinese Teaching (End of Semester)

Dear Parent,

The semester is coming to an end. Please take 5 minutes to reflect on this semester's AI-assisted Chinese lessons and answer the following questions.

1. Have you observed your child mentioning the Al or showing changes in Chinese learning at home? (Please describe specifically)

2. What is your current attitude toward using AI in Chinese class?

☐ Supportive, hope it continues next semester
☐ Conditionally supportive (please specify conditions: _____)
☐ Neutral
☐ Opposed, hope it stops next semester
☐ Uncertain

3. Compared to the beginning of the semester, has your attitude changed? Why?

4. If you have any suggestions for the teacher or the school, please write them here:

Thank you for your support throughout the semester!

Appendix C: Classroom Observation Record Forms (Blank Templates)

Meaning Negotiation Event Record Form (Blank)

Time	Student	Context Description	AI's Response	Student's Response	Negotiation Process Description	Outcome & Notes

Completion Instructions:

- **Time:** Approximate time after the activity begins (e.g., "approximately 3 minutes after activity starts")
- **Student:** Use pseudonym (e.g., S3, S7)
- **Context Description:** What task was the student doing? What did they say?
- **AI's Response:** What did the AI say? (verbatim record)
- **Student's Response:** What did the student say? What gestures did they make? What was their facial expression?
- **Negotiation Process Description:** Describe the entire negotiation process in narrative form (what happened? What strategies did the student try?)
- **Outcome:** Did the AI eventually understand? Did the student give up? Did they seek help?
- **Observer's Notes:** Researcher's preliminary analysis and reflections

AI Role Perception Record Form (Blank)

Time	Student	Observational Evidence (speech, behavior, affective expression)	AI's Current Role (and basis for judgment)	Role Changed?	Observer's Notes

Completion Instructions:

- **Time:** Time period within the activity (e.g., "first 5 minutes of activity," "middle of activity")

- **Student:** Use pseudonym
- **Observational Evidence:** What did the student say? What did they do? What was their facial expression? What did they say to peers?
- **Al's Current Role:** Tool / Peer in Need / Fallible Opponent / Coach / Teacher / Other
- **Basis for judgment:** Why do you think the student assigned this role to the AI?
- **Role Changed?:** Has it changed compared to before? What triggered the change?
- **Observer's Notes:** Researcher's preliminary analysis and reflections

Note: All student names are pseudonyms. Some example responses are composites for illustration.

Appendix D: Student Task Card Templates

Three student task card templates are provided below, corresponding to Week 1 ("My Family"), Week 5 ("Asking for Directions and Transportation"), and Week 13 ("My Weekend"). Each task card includes:

- Main task card body (printable for distribution to students)
- Completion instructions (for teacher reference)

All task cards are designed to A4 size and can be photocopied directly. Language uses simple Chinese + English keywords, suitable for HSK Level 3 students.

Task Card 1: Week 1 - "My Family"

Student Use: Chinese Class Task Card – My Family
Name (Code): _____ Date: _____ Partner: _____

Part 1: Write (Before speaking to the AI)

Please write three Chinese sentences describing your family members. Use at least one word from the box below for each sentence. *Words you can use: tall / short / fat / thin / funny / serious / more...than... / a little*
Sentence 1 (describe one person): Sentence 2 (compare two people): Sentence 3 (describe personality):

Part 2: Speak with the AI

Open ChatGPT and say your three sentences to the AI (the "Little Silly Robot"). Did the Al understand?

- **Sentence 1:** ☐ Understood first time ☐ Understood after __ tries ☐ Never understood
- **Sentence 2:** ☐ Understood first time ☐ Understood after __ tries ☐ Never understood
- **Sentence 3:** ☐ Understood first time ☐ Understood after __ tries ☐ Never understood

If the AI didn't understand, what did you do? (Select all that apply)

☐ Repeated it (louder/slower)
☐ Changed the wording
☐ Used English

☐ Used gestures
☐ Asked a peer
☐ Other: _____

Part 3: Teach and draw
The teacher will have the AI draw your family members. Your group chooses one sentence and writes it below. The teacher will input this sentence into the AI to see if the AI draws it correctly.

- Our group's chosen sentence: _____
- Did the AI draw correctly? ☐ Yes ☐ No
- If not, what should the correct sentence be? _____

Part 4: Self-assessment

☐ Today when I spoke with the AI, I felt: ☐ Nervous ☐ Relaxed ☐ Fun ☐ Bored ☐ Other: _____
☐ I think the AI is like: ☐ A robot ☐ A little kid ☐ A teacher
☐ My favorite part today was: _____

Teacher's Completion Instructions (Week 1)

Item	Instructions
Time allocation	Part 1: 5 min, Part 2: 10 min, Part 3: 5 min, Part 4: 2 min
Grouping	Pairs; one speaks to AI, the other records (switch roles)
AI setting	Use "Little Silly Robot" role (intentionally makes mistakes)
Key reminders	Remind students not to type real names; tell students "The AI sometimes doesn't understand; that's okay, you can teach it"
Task card collection	Collect after class for analysis of students' output modification and affective feedback

Task Card 2: Week 5 - "Asking for Directions and Transportation"

Student Use: Chinese Class Task Card - Giving Directions to Directionally Challenged Xiao A
Name (Code): _____ Date: _____ Partner: _____

Our Map
[Park] ↑ [School] → [Supermarket]
↓ [Subway Station]
Legend:

- From School to Park: Go straight first, then turn left (3 minutes)

- From School to Supermarket: Go straight, turn right at the second intersection (5 minutes)
- From School to Subway Station: First turn left, then go straight, you'll see the traffic light (4 minutes)

Part 1: Choose a destination (Before speaking to the AI)

- I choose to go to: ☐ Park ☐ Supermarket ☐ Subway Station
- My route (write in Chinese): _____

Part 2: Speak with the AI

Open ChatGPT and say your route to the AI ("Directionally Challenged Xiao A"). The Al will intentionally make mistakes. You need to correct it.

- Have I finished saying my route? ☐ Yes ☐ No
- The Al's first mistake (what wrong thing did the AI say?): _____
- How did I correct it? (write what you said): _____
- Did the Al understand?
  ☐ Yes ☐ No
- If the AI still didn't understand, what did I do? (Select all that apply)

  ☐ Said it again
  ☐ Changed the wording
  ☐ Used hand gestures for directions
  ☐ Asked a peer
  ☐ Other: _____

Part 3: Challenge level (Optional)

After completing one route, try a second route.

- My second chosen destination: ☐ Park ☐ Supermarket ☐ Subway Station
- My route: _____
- How many times did the Al make mistakes? ☐ 1 time ☐ 2 times ☐ 3 or more times
- Did I say anything about the Al to my peer?
  ☐ Yes (what did I say? _____) ☐ No

Part 4: Self-assessment

- When I corrected the AI today, I felt: ☐ Very annoyed ☐ Okay ☐ Very fun ☐ Other: _____
- I think "Directionally Challenged Xiao A" is like: ☐ A lost person ☐ A little kid ☐ A game opponent ☐ Other: _____
- My favorite part today was: _____

Teacher's Completion Instructions (Week 5)

Item	Instructions
Time allocation	Part 1: 5 min, Part 2: 12 min, Part 3: 8 min (optional), Part 4: 2 min
Grouping	Pairs, taking turns as "direction giver" and "recorder"
AI setting	Use "Directionally Challenged Xiao A" role (intentionally reverses directions)
Key reminders	Remind students "The AI will mix up left and right; be ready to correct it"; encourage students to use gestures
Task card collection	Collect after class; focus on descriptions of the "correction process" and affective feedback

Task Card 3: Week 13 - "My Weekend"

Student Use: Chinese Class Task Card – My Weekend Timeline
Name (Code): _____ Date: _____

Part 1: Write (Handwritten draft - do NOT ask the Al first)
Please write three things you did last Saturday. Use "First... then... finally..."

- First (first thing): _____
- Then (second thing): _____
- Finally (third thing): _____
- Complete sentence (connect the three things above): _____

Part 2: Turn it into a timeline with the AI
Open ChatGPT and enter your sentence. Tell the AI: "Please turn my weekend into a timeline."

- The timeline the AI generated (copy it below): _____
- Is the Al's order correct? ☐ Yes ☐ No
- If not, what was wrong? _____
- How did I correct it? (write what you said): _____
- Did the AI correct it? ☐ Yes ☐ No

Part 3: Turn it into a story with the AI (Challenge level)
Say to the AI: "Please turn my timeline into a story."

- The story the AI generated (copy it below): _____
- Does this story sound like yours? ☐ Yes ☐ Not really
- What would you change? _____

Part 4: Self-assessment

- Did the AI understand your sentence today? ☐ Understood first time ☐ Understood after __ tries
- I think the AI is like: ☐ A tool ☐ An assistant ☐ A silly classmate ☐ Other: _____
- Did you "teach" the AI today?
  ☐ Yes (what did you teach? _____) ☐ No

- Was the process of "teaching" the AI helpful to you? ☐ Helpful ☐ Not helpful
- Why? _____

Teacher's Completion Instructions (Week 13)

Item	Instructions
Time allocation	Part 1: 5 min, Part 2: 10 min, Part 3: 8 min (optional), Part 4: 2 min
Grouping	Can work individually or in pairs to check each other's work
Al setting	Use normal AI (no intentional mistakes), but students need to actively check whether the Al's output is correct
Key reminders	Emphasize "write your draft first, then ask the AI" - do NOT let the AI write it directly; remind students to check whether the Al's order is correct
Task card collection	Focus on comparison between "handwritten draft" and "AI output," and students' reflections on "teaching the AI"

General Completion Instructions (For Teacher Reference)

Principles for Using Task Cards

Principle	Explanation
Handwriting first	Students must write by hand first, then converse with the AI. This ensures students have thought about their own output rather than directly copying the Al's responses.
Immediate recording	During the AI conversation, encourage students to make simple marks on the task card (checkmarks, keywords) while speaking. They can complete more details after class.
No grading	Task cards are not used for grading, only for formative feedback and data collection for this study. Teachers may write brief comments on task cards (e.g., "Good job using gestures!").
Collection and feedback	Collect task cards after class; teacher quickly reviews them to understand students' difficulties (e.g., many students stuck on the same word) and addresses them in the next lesson.
Privacy protection	Use only codes on task cards, not real names. Store scanned copies securely; destroy paper versions after scanning.

Handling Common Problems with Task Cards

Problem	Teacher Response
Student directly copies AI's response instead of writing their own	Remind them to "write your own sentence first, then speak to the AI." If it happens repeatedly, have an individual conversation.
Student does not complete the "Self-assessment" section or responses are too brief (e.g., only "good," "not good")	Leave 2 minutes at the end of the activity specifically for filling this out; remind students while circulating. Combine checkboxes with open-ended questions on the task card to reduce difficulty.
Student loses the task card	Prepare 5-10 blank task cards as backups for each lesson; encourage students to keep them in a dedicated folder.

Appendix E: Student Semi-Structured Interview Protocol

Instructions for Use This interview protocol is designed for one-on-one semi-structured interviews with focal students at Week 6 (mid-semester) and Week 18 (end of semester). Each interview lasts 20-25 minutes.

Preparation before the interview:

- Choose a quiet, distraction-free location (e.g., hallway outside the classroom, empty small meeting room)
- Prepare recording equipment (test in advance)
- Prepare a printed copy of the interview protocol (can be referred to, but do not read verbatim)
- Prepare water and tissues

Ethical reminders for the interview:

- Confirm the student's consent to audio recording before starting the interview
- Inform the student that they may skip any question or end the interview at any time
- Inform the student that this is not a test, there are no right or wrong answers, and it will not affect their grade
- Use language the student can understand (Chinese + English keywords)

Interview Opening Script (Verbatim) "*Note: All student names referenced in this appendix (e.g., S3, S7, S11) are pseudonyms in the S1-S14 format used throughout the manuscript.*

Thank you for being willing to talk with me. This is not a test today, and there are no right or wrong answers. Your responses will not affect your Chinese class grade. I want to know what you really think whether you like the Al or don't like it, you can say it directly.

I will record the audio, just so I don't misremember things when I write them down later. Only I can hear the recording. I won't use your real name in the thesis, and no one else will know it was you who said these things.

If you don't want to answer a question, you can say 'I don't want to say' at any time.

Are you ready? Then let's begin."

Part 1: Experience with Al Dialogue

(Corresponding to RQ1: Meaning negotiation and output)

Question 1.1

"In class, we used ChatGPT (that 'Little Silly Robot' or 'Directionally Challenged Xiao A'). Do you remember what you said to it? Can you give an example?"
Probes (use flexibly based on response):

- "Did the AI understand you at that time? If not, what did you do?"
- "What exactly did you say? Can you say it to me again?"

Question 1.2

"The AI sometimes 'doesn't understand' or 'makes mistakes.' Do you remember any mistakes it made? What did you do then?"
Probes:

- "Did you say it again, or did you change the wording?"
- "Did you use gestures? Or use English?"
- "Did you look at a peer? Or say something to a peer?"

Question 1.3

"If the sentence you said wasn't quite right, the AI would ask you to say it again. How do you usually change it in that situation? Can you give an example?"
Probes:

- "Do you change one word, or rephrase the whole sentence?"
- "Did the AI understand after you changed it? How did you feel?"

Question 1.4 (Mid-semester interview) / Question 1.4 (End-semester interview variation)

Mid-semester version:
"Since the beginning of the semester, have you noticed any changes in how you speak to the AI? For example, before you had to say it several times for the AI to understand, but now it understands the first time?"
End-semester version:
"Over this semester, do you feel you have improved in your ability to speak to the AI? Can you give an example?"
Probes:

- "What do you think made you improve (or not improve)?"

- "Have you noticed that before speaking to the Al, you think through the sentence in your head first?"

Part 2: Feelings about the AI (Anxiety and Affect)

(Corresponding to the Affective Filter Hypothesis)

Question 2.1

"In Chinese class, when do you feel nervous or afraid to speak?"
Probes:

- "Do you feel nervous when speaking to the AI? What about speaking to classmates? What about speaking to the teacher?"
- "If you had to rank them, which makes you most nervous?"

Question 2.2

"When speaking Chinese to the AI, what happens if you make a mistake? Do you feel uncomfortable inside?"
Probes:

- "If the AI says 'I don't understand,' do you feel it's the Al's problem or your problem?"
- "If a classmate hears the Al not understanding you, do you feel embarrassed?"

Question 2.3 (End-semester interview)

"Over this semester, has speaking to the AI made you more willing to speak Chinese in real conversations? Can you give an example?"

Part 3: What the AI Is Like (Role Perception)

(Corresponding to RQ2: AI role perception)

Question 3.1

"If you had to describe the 'Little Silly Robot' or 'Directionally Challenged Xiao A' in one sentence, what do you think it is like? Is it a teacher, a classmate, a game, a dictionary, or something else?"
Probes:

- "Why do you think it is like [student's answer]?"
- "Does it sometimes seem like this, and sometimes like something else?"

Question 3.2

"When you hear the AI say 'I don't understand,' do you feel like you are teaching it, or like it is testing you?"
Probes:

- "Do you prefer it to say 'I understand' or 'I don't understand'? Why?"

Question 3.3 (End-semester interview)
"At the beginning of this semester, how did you feel about the AI? How do you feel now? Has there been any change?"
Probes:

- "If you could design an AI to help you learn Chinese, what would you want it to be able to do?"

Question 3.4 (Optional, use if time permits)
"Do you think the AI could replace a Chinese teacher? Why or why not?"

Part 4: Privacy and Safety

(Corresponding to RQ3: Privacy and ethics)
Question 4.1
"Do you know that the Al we use in class records what you say? Does that worry you?"
Probes:

- "The teacher clears the records after each class. Did you know that? Do you think it's important?"
- "If the AI remembered every word you said, would you still say the same things to it?"

Question 4.2 (Optional)
"Did you tell your parents at home that we use AI in class? What did they say?"

Part 5: Summary and Suggestions

Question 5.1
"If you could give one piece of advice to next semester's Chinese class about how to use Al better, what would you say?"
Question 5.2
"Is there anything else you want to tell me? Anything about AI or Chinese class?"
Interview Closing Script
"Thank you! What you've said is very helpful to me. If you think of anything else later, you can come find me anytime. Goodbye!"

Appendix: Interview Questions and Corresponding Theoretical Framework

Question Number	Core Content	Corresponding Theory/Concept
1.1-1.2	Specific experiences with Al dialogue	Interaction Hypothesis (negotiation of meaning)
1.3	Methods of output modification	Output Hypothesis (hypothesis testing)
1.4	Self-perceived changes	Output Hypothesis (metalinguistic reflection)
2.1-2.2	Anxiety and affective experience	Affective Filter Hypothesis
2.3	Al's impact on anxiety	Affective Filter Hypothesis
3.1-3.2	Perception of Al roles	Al role framework
3.3	Changes in role perception	Role fluidity
3.4	AI vs. teacher	Al role framework
4.1-4.2	Privacy awareness	RRI/Privacy framework

Interview Technique Reminders

Not Recommended	Recommended Alternative
"Don't you think the Al is very helpful?" (leading question)	"Do you find the Al helpful? Why or why not?"
"What do you think of the AI?" (too abstract)	"Do you remember last time when the AI said 'I don't understand'?"
"Is the AI like a teacher or like a friend? Do you find it useful?" (double question)	First ask "What is the AI like?" then ask "Do you find it useful?"
"That's great!" (evaluative response)	"Mm-hmm, I hear you." (neutral response)

Appendix F: Parent Semi-Structured Interview Protocol

Instructions
This interview protocol is designed for one-on-one semi-structured interviews with parents of focal students, conducted in Week 2 (beginning of semester) and Week 18 (end of semester). Each interview lasts 15-20 minutes.

Pre-interview preparation:

- Schedule an appointment with parents in advance, confirm whether online (e.g., Zoom/WeChat voice) or in-person
- Prepare recording device (test in advance; for online interviews, use recording software)
- Print a copy of the interview protocol (can refer to it, but avoid reading verbatim)
- Prepare a copy of the parent letter for reference during the interview

Ethical considerations:

- Confirm parents' consent to audio recording before starting the interview
- Inform parents that they may skip any question or end the interview at any time
- Inform parents that all information will be anonymized, and no identifiable information will appear in the paper
- Respect parents' language preference (interviews can be conducted in English or Chinese)

Interview Opening Script (verbatim)
"*Note: Student identifiers (e.g., S3, S7, S11) follow the S1-S14 format used throughout the manuscript.*

Thank you for taking the time to talk with me. I am [name], your child's Chinese teacher. I am conducting a study on 'AI-Assisted Chinese Learning' and would like to hear your genuine thoughts whether you support or have concerns, both are very helpful to me.

This conversation will be recorded, but only for research analysis. Your real name will not be used in the paper, nor will your child's name. You may choose not to answer any question at any time.

Shall we begin?"

Part 1: Beginning of Semester Interview (Week 2)

1.1 Initial Reaction to the Parent Letter
Question P1.1
"Do you remember the 'Letter to Parents' I sent at the beginning of the semester? What was your first reaction after reading it?"
Follow-up probes:

- "What concerned you the most at that time?"

- "What did you agree with the most?"
- "Did you discuss this letter with your child?"

Question P1.2
"In your impression, do you think introducing Al into the Chinese classroom is mainly a good thing, a bad thing, or mixed? Why?"

1.2 Specific Concerns
Question P1.3
"Regarding Al data privacy, what concerns you the most? For example: your child's voice being recorded, conversations being saved, information being sold to third parties, or being used for other purposes?"
Follow-up probes (show card for parents to choose/rank):
"Which of the following is your biggest concern?"

- A. Child's real name or photo being recorded by AI
- B. Child's voice being saved
- C. What the child says being seen by third parties
- D. AI pushing inappropriate content to the child
- E. Child spending too much time on screens
- F. Other: _____

Question P1.4
"Apart from privacy, do you have any other concerns? For example, screen time, technology dependence, or reduced interaction with real people?"
Follow-up probes:

- "If so, which one concerns you the most?"
- "Do you think one AI lesson per week (about 50 minutes) is acceptable? Why?"

1.3 Expectations for Child's Chinese Learning
Question P1.5
"What are your expectations for your child's Chinese learning this semester? In which areas do you hope to see improvement?"
Follow-up probes:

- "Do you think AI might help in these areas? Why?"

Question P1.6
"If you could give me one piece of advice about using Al, what would it be?"

Part 2: End of Semester Interview (Week 18)

2.1 Observed Changes
Question P2.1
"Over this semester, have you noticed your child mentioning Chinese class or Al at home? Do you have any specific examples or stories?"
Follow-up probes:

- "Did your child voluntarily speak Chinese at home (e.g., giving you directions, introducing family members, teaching you Chinese words)?"
- "Did your child complain about or dislike anything?"

Question P2.2

"Do you think your child's interest in learning Chinese has changed this semester? If so, what role do you think Al played?"
Follow-up probes:

- "Did your child voluntarily use AI at home (e.g., looking up words, practicing conversation)?"
- "If so, what did you observe about how they used it?"

2.2 Change in Attitude

Question P2.3

"Now that the semester has passed, has your attitude toward using Al in the Chinese classroom changed compared to the beginning of the semester?"
Follow-up probes:

- "If you had to choose now: continue using Al next semester / reduce usage / stop completely - which would you choose? Why?"

Question P2.4

"What about the concerns you had at the beginning of the semester? How are they now? [Follow up based on concerns mentioned in the first interview]"
Follow-up probes:

- "If you still have concerns, what are they mainly?"
- "If your concerns have lessened, what made you feel less worried?"

2.3 Re-evaluation of AI's Value

Question P2.5

"Do you think Al has helped your child's Chinese learning? If so, what is the biggest help?"
Follow-up probes:

- "Is there anything you feel 'could not be done without AI'?"
- "Is there anything you feel 'AI is not good for'?"

Question P2.6

"Are there any things you think Al should 'absolutely not do'? For example: evaluating your child's personality, giving rankings, replacing the teacher's emotional support?"

2.4 Suggestions for the Future

Question P2.7

"If you could write one piece of advice about Al use to the school or the teacher, what would it be?"
Follow-up probes:

- "If AI continues to be used next semester, what would you like the teacher to do differently?"

- "What information would you like parents to receive?"

Question P2.8

"Is there anything else you would like to tell me? About AI, Chinese class, or your child's learning anything is fine."

Interview Closing Script

"Thank you very much for sharing your thoughts openly. Your feedback is very important for improving my teaching. If you think of anything else later, please feel free to email me anytime. Thank you!"

Appendix: Mapping Interview Questions to Theoretical Framework

Question Number	Core Content	Corresponding Theory/Concept
P1.1-P1.2	Initial reaction to parent letter	Informed consent, transparency
P1.3	Specific privacy concerns	OECD Privacy Principles
P1.4	Screen time, technology dependence concerns	Parent concern categories (Akgun & Greenhow, 2022)
P1.5-P1.6	Expectations for Chinese learning	Technology Acceptance Model (perceived usefulness)
P2.1-P2.2	Observed changes in child	External validation (triangulation with classroom observation)
P2.3-P2.4	Change in attitude	RRI framework (responsiveness, reflexivity)
P2.5	Re-evaluation of Al's value	Technology Acceptance Model
P2.6	Boundaries of what AI "should not do"	RRI framework (ethics)
P2.7-P2.8	Suggestions for future use	Inclusivity, continuous improvement

Parent Concerns Quick Reference Card

(Can be shown to parents during the interview to help them articulate their concerns)

Interview Record Form Example

(For researcher's quick on-site notes)

Concern Type	Specific Content	Mitigation Strategies in This Study
Data Privacy	Child's conversations recorded, voice saved	No real names, conversations cleared after class
Screen Time	Child spends too much time looking at screens	Only 1 AI lesson per week; AI use about 20 minutes per lesson
Technology Dependence	Child relies on Al instead of thinking independently	Task cards require "write first, then ask AI"
Content Appropriateness	AI generates inappropriate content	Teacher monitoring, content filtering enabled
Social-Emotional Development	Child has less interaction with real people	Three traditional lessons per week; peer interaction also in Al lessons

Parent Code	Interview Time	Main Concerns (Beginning)	Attitude Change (End)	Key Quotes
P3	Week 2 / Week 18			
P7	Week 2 / Week 18			
P11	Week 2 / Week 18			

Appendix G: Teacher Reflection Log Template

Instructions

This log template is for the researcher (teacher-researcher) to use after each weekly lesson, for recording observations, decisions, unexpected events, ethical dilemmas, and personal reflections during the teaching process. Each entry should be approximately 500-1000 words, written in narrative form rather than as a checklist.

Time of writing: Complete on the same day as each Al lesson (preferably within 30 minutes after class, while memory is fresh)

Writing principles:

- Write in narrative form; do not just list bullet points
- Record both "what happened" and "what I was thinking at the time"
- Include both successes and failures, smooth moments and difficulties
- Record emotional responses ("I felt...")
- Record follow-up action plans ("Next time I will...")

Log Template Structure

Section	Content	Suggested Length
Basic Information	Week number, date, lesson topic, Al tools	Quick fill
Lesson Overview	What happened? How did it feel overall?	100-150 words
Technical Operation	Did AI tools work properly? Any malfunctions?	50-100 words
Student Observations	How did focal students perform? Any unexpected behaviors? Who was engaged? Who struggled?	150-250 words
Teaching Decision Reflection	What adjustments did I make? Why?	100-150 words
Ethical Dilemma Record	Any privacy or ethical issues? How were they handled?	100-150 words
Emotion and Reflexivity	How did I feel? How might this affect my judgment?	50-100 words
Next Week Action Plan	What will I do differently next week based on today's experience?	50-100 words

Blank Log Template

Teacher Reflection Log
Week: _____
Date: _____
Lesson Topic: _____
AI Tools Used: ☐ ChatGPT ☐ Duolingo ☐ Canva ☐ Wordwall ☐ Other: _____

I. Lesson Overview
(What happened? How did it feel overall?)

II. Technical Operation
(Did Al tools work properly? Any malfunctions? Did students encounter technical problems?)

III. Student Observations
(How did focal students perform? Any unexpected or interesting behaviors? Who was particularly engaged? Who struggled?)

IV. Teaching Decision Reflection
(What adjustments did I make today? Why? What was the original plan?)

V. Ethical Dilemma Record
(Did any privacy or ethical issues arise? If so, how did I handle them?)

VI. Emotion and Reflexivity
(How did I feel during class today? Excited? Frustrated? Anxious? How might these feelings affect how I observe and record?)

VII. Next Week Action Plan
(Based on today's experience, what will I do differently next week?)

Completed Example (Week 1)

The following is a completed example for reference.

Teacher Reflection Log
Week: 1
Date: September 5, 2025
Lesson Topic: "My Family" - First AI Conversation
AI Tools Used: ☒ ChatGPT ☐ Duolingo ☐ Canva ☐ Wordwall ☐ Other:

I. Lesson Overview
Today was the first time students used AI conversation. Overall, it was a bit more chaotic than expected technical preparation took longer than planned, and two students' tablets needed to be logged in again. The activity itself went relatively smoothly, and most students completed the three sentences on their task cards. However, S7 (Korean L1, lower proficiency) was clearly very nervous, speaking very softly. When the AI said "I don't understand," he froze. S11 (French L1, high proficiency) was very calm but seemed slightly impatient with the Al's "stupidity." S3 (English L1, intermediate-high proficiency) performed best, patiently explaining "traffic light." Overall, the goal of the first lesson-to get students exposed to Al-was achieved. But students clearly have not yet adapted to the fact that the AI can "make mistakes."

II. Technical Operation

- ChatGPT voice function worked normally, but there was a 2-3 second delay, requiring students to wait.
- One tablet (S14) had not updated ChatGPT and could not open it. I had him work with S13.
- Screen recording worked normally; conversations for S3, S7, and S11 were all captured.
- After class, I reminded the whole class to close their ChatGPT windows. Most students did so. S12's window was still open; I helped him close it.

III. Student Observations
S7: Very nervous. When he said "My father is tall" to the AI, his voice was so soft that I could barely hear it standing next to him. After the AI said "I don't understand," he was silent for about 5 seconds, then looked at his partner. His partner

mouthed "tall," and he turned back and said "Tall means... tall." He did not smile or show any sign of relaxation. After the activity, he quickly closed his tablet, as if relieved.

S3: Relatively calm. When the AI did not understand "traffic light," she did not get nervous. She paused, thought for a moment, then explained using a mix of Chinese and English. She used gestures (pointing to the intersection outside the window). Her expression was focused, not frustrated. She seemed to treat "teaching the AI" as a natural thing.

S11: Very calm, but slightly impatient. Her sentences were accurate, and the AI understood them immediately. However, when the AI made a mistake during confirmation check (saying "McDonald's" instead of "park"), she sighed and corrected the AI with slightly faster speech. She did not discuss the Al with her partner or show any excitement.

Other observations: Most students showed surprise when they first heard the AI say "I don't understand"-some paused, some frowned, some turned to look at their partners. No student showed obvious excitement or enjoyment. This reminds me: the "stupid AI" setting might increase anxiety in the beginning rather than reduce it. I need to remind students next week, "This AI is a bit stupid; you need to teach it."

IV. Teaching Decision Reflection

Today's activity took about 5 minutes longer than planned because of the time spent on technical preparation. Instead of cutting students' conversation time, I shortened the final "presentation" segment. I think this was the right adjustment because conversation practice is more important than presentation. One decision I am still thinking about: When S7 got stuck, I walked over wanting to help him, but he had already started asking his partner for help. I did not intervene and let his partner help him. I think this was the right decision-peer support is also learning. But if he had not had a partner, I might have needed to intervene.

V. Ethical Dilemma Record

Today, S8 accidentally typed his full name during the AI conversation. I walked over, helped him delete that conversation, and gently reminded him, "Don't type your name. Use your code name." He nodded. This incident reminds me: one reminder is not enough. Students are used to typing their names into various apps and can easily forget. Starting next week, I will give a 10-second reminder at the beginning of each Al lesson: "Don't type your name. Use your code name." Also, one student asked me today, "Will the Al remember what I said?" I said no, because we close the window after class. This made me realize that students themselves have privacy awareness. I should more explicitly tell them why we close the window.

VI. Emotion and Reflexivity

To be honest, I was a bit nervous before class today this was my first time using Al for a formal activity in a real classroom. I was afraid of technical problems and afraid students wouldn't like it. During class, when I saw how nervous S7 was, I felt a bit anxious and wanted to help him. But I held back because I knew that if I intervened every time a problem arose, students would never learn to solve problems on their own. After watching the recording, I noticed that I paid more attention to S7 than to S11. This might be because S7's problems were more "visible"-his nervousness,

his silence. I need to remind myself not to focus only on students who are "having problems" but also on those who are "not having problems," because their calmness might also carry important meaning.

VII. Next Week Action Plan

1. Give a 10-second reminder at the beginning of each AI lesson: "Don't type your name. Use your code name."
2. Before starting the activity, spend 1 minute telling students: "This AI sometimes doesn't understand. That's normal. It's not your fault. You need to teach it."
3. Pay more attention to S11's behavior next week is her calmness because she doesn't need AI at all, or just because she doesn't show it?
4. Consider giving S7 some encouragement. He was very nervous today, but if he can overcome that, he might become the biggest beneficiary of the AI.

Completed Example (Week 5)

The following is a mid-semester completed example, showing how the log becomes "thicker" as the semester progresses.

Teacher Reflection Log

Week: 5

Date: October 15, 2025

Lesson Topic: "Asking for Directions" - Giving Directions to Directionally Challenged Little A

AI Tools Used: ☒ ChatGPT ☐ Duolingo ☐ Canva ☐ Wordwall ☐ Other:

I. Lesson Overview

Today's lesson went very smoothly, even better than expected. Students were visibly much more relaxed than in Week 1, and there was a lot of laughter. The AI setting of "intentionally getting directions wrong" worked very well-students quickly figured out the Al's pattern and started actively correcting it. S7 today was like a different person: from speaking softly in Week 1 to loudly correcting and saying to his partner "It's so stupid"-the change was dramatic. However, some students (like S11) did not show excitement, only calmly corrected. This reminds me that the "gamified" Al setting is not effective for everyone. Some students just treat it as a tool, and that's fine.

II. Technical Operation

ChatGPT worked normally today with no delays. Screen recording also worked normally. One group (Group 5) had low battery on their tablet; I had them move to another spot.

III. Student Observations

S7: The biggest surprise today. His voice was still a bit soft at the beginning, but when the Al first got the direction wrong, he immediately corrected loudly: "No! Not left turn. Go straight first!" Then he turned to his partner, smiled, and said "It's

so stupid." After that, his whole demeanor changed-leaning forward, exaggerated gestures, firm voice. Even when it wasn't his turn to speak, he leaned over to look at the Al's responses to his partner. He really treated this as a game.
S3: Still the same patient, calm. When the AI made a mistake, she was neither excited nor frustrated, just calmly corrected. She used gestures to show directions, saying "Look, I'll show you with my hand." Her posture was like a teacher teaching a student, not a player playing a game.
S11: Calm, efficient. Her sentences were accurate, and the AI understood immediately. But when the AI made a mistake during confirmation check, she sighed and said, "No, at McDonald's it's left turn, to the park it's right turn." She didn't smile or say anything to her partner. She was just completing the task.
Other observations: A lot of laughter today. Every time the AI gave an obviously wrong direction, several groups laughed or exclaimed simultaneously. One girl said to her partner, "Is it not awake today?"-this suggests that students no longer see the Al's mistakes as "their failure," but as a "feature" of the AI itself.

IV. Teaching Decision Reflection

There was one change in today's activity design: I did not have the AI "intentionally go off-topic," only "get directions wrong." This setting was simpler and more predictable. Based on student reactions, this was the right decision-because students quickly figured out the pattern and started actively correcting. If the Al's mistakes had been too random, students might have felt frustrated rather than excited. I'm still thinking: should I make the Al's mistakes more challenging? For example, have the AI change "go straight first, then turn right" to "turn right first, then go straight"? But today's results were already very good, and I don't want to "fix something that isn't broken."

V. Ethical Dilemma Record

No privacy issues occurred today. I gave the "don't type your name" reminder at the beginning of the activity, and students remembered. One small issue: when S7 said "It's so stupid," he was smiling and meant no harm. But if other students imitate this, they might say it more harshly. I need to pay attention to this. If students start "insulting" the AI, I may need to intervene and remind them, "The AI isn't really stupid; the teacher made it that way." Also, one student (S13) showed mild frustration after the AI made several mistakes in a row. She didn't laugh like S7; she sighed. I walked over and asked, "What's wrong?" She said, "It keeps getting it wrong." I said, "It's supposed to get things wrong. You need to teach it." She nodded and continued. But later I thought: should I have given her more help? Or let her solve it herself? I chose to let her solve it herself because her frustration was mild. But if a student becomes very frustrated in the future, I may need to intervene directly.

VI. Emotion and Reflexivity

I was very happy during class today. Seeing S7 go from nervous in Week 1 to excited today made me feel that this curriculum design really works. But I also realize that my "happiness" might affect my observations will I unconsciously focus only on S7's "successes" and ignore his "difficulties"? I need to remind myself to record his frustrating moments as well, not just the highlights. Also, I may not have observed

S11 enough today. She was very calm, but I didn't pay much attention to her. Does she find this activity too easy? Is she bored? I need to pay more attention to her next week.

VII. Next Week Action Plan

1. Pay more attention to S11 next week to see if she really finds the activity too easy.
2. Consider designing a "challenge level" for high-proficiency students-for example, having the AI make more complex mistakes.
3. Continue the "don't type your name" reminder before each lesson.
4. Document S7's trajectory of change his excitement today: temporary or sustainable?

Log Writing Prompt Card

If unsure what to write	**Ask yourself these questions**
Lesson Overall	What surprised me most today? What went most smoothly? What was most difficult?
Technology	Did the AI tools malfunction? How did I resolve it? How did students react?
Students	Who stood out today? Who struggled? Who surprised me?
Teaching Decisions	What unplanned adjustments did I make? Why? What was the effect?
Ethics	Did any students type personal information? Did the AI generate inappropriate content? How did I handle it?
Emotion	How did I feel today? How might this affect my judgment?
Next Time	If I taught this lesson again next week, what would I do differently?

Appendix H: Coding Scheme Example

Instructions
This appendix presents the qualitative coding scheme used in this study. The coding scheme is the core tool of thematic analysis, used to label and categorize units of meaning in raw data (classroom observation records, interview transcripts, reflection logs, etc.).

This coding scheme adopts a mixed strategy: some codes are pre-established based on theoretical frameworks (theory-driven), and some codes emerge from the data (data-driven). The codes are organized into three levels:

- **Level 1 codes:** Descriptive codes that directly label data content
- **Level 2 codes:** Interpretive codes that provide preliminary categorization of Level 1 codes
- **Level 3 codes:** Thematic codes that form the core themes of the research findings

This appendix presents the final version of the coding scheme, including code names, definitions, inclusion criteria, exclusion criteria, and data examples.

H.1 Overview of Coding Structure

Table H.1 Overall coding structure:

Code Category	Number of Level 1 Codes	Number of Level 2 Codes	Research Question
Negotiation of Meaning Behaviors	8	3	RQ1
Output Modification	5	2	RQ1
Affective Expression	6	3	RQ1/RQ2
AI Role Perception	7	4	RQ2
Privacy and Ethics	5	2	RQ3
Teacher Reflection	4	2	RQ3

H.2 Detailed Code Definitions

H.2.1 Negotiation of Meaning Behavior Codes

These codes are used to label interactive adjustments made by students to overcome comprehension difficulties with the AI.

Table 7 Table H.2.1 Coding scheme for negotiation of meaning behavior

Code Name	Definition	Inclusion Criteria	Exclusion Criteria	Data Example
negotiation_repetition	Student repeats the original utterance in identical or highly similar form, usually accompanied by increased volume, slower speech, or changed intonation	Student repeats utterance with essentially same words and word order; purpose is to make AI understand	Student changes words or sentence structure (code as negotiation_paraphrase)	"Turn right! Right! Not left!" (S7, Week 5)
negotiation_paraphrase	Student expresses the same meaning using different words, involving lexical substitution, word order adjustment, or structural changes	Student uses different words or structure from original utterance; core meaning remains unchanged	Student simply repeats original utterance (code as negotiation_repetition); student changes meaning	"Traffic light means... traffic light. The thing at the intersection, red light stop green light go." (S3, Week 1)
negotiation_simplification	Student reduces number of words or shortens sentence structure for more direct expression	Student's output is shorter than original utterance with fewer words; core information retained	Student simply repeats original utterance; student paraphrases but length is unchanged or longer	"Then... then... then traffic light." (S7, Week 5)
negotiation_code_switching	Student inserts English words into Chinese expression	Student uses English word or phrase; English is used to replace or supplement Chinese expression	Student uses other language (e.g., Korean, French); student uses complete English sentence	"Tall means... tall." (S7, Week 1)
negotiation_gesture	Student uses gestures, pointing, body movements, or other non-verbal means to support expression	Student makes observable non-verbal action; action occurs simultaneously with or immediately after verbal expression	Student only speaks with no observable action	S7 uses hand gesture pushing forward (Week 5); S3 uses fingers to count three items (Week 13)
negotiation_peer_help	Student turns to peer for help, or peer voluntarily offers assistance	Student has verbal or non-verbal interaction with peer; content is related to A conversation task	Student discusses topics unrelated to task with peer	S7 turns to peer and asks "How do you say 'traffic light'?" (Week 5)
negotiation_abandonment	Student gives up on current sentence or task without further attempts	Student stops speaking, closes tablet, or says "Forget it" or "I don't know how to say it"	Student pauses briefly then continues trying (code under other categories)	"Forget it" after three consecutive attempts and moves to next task (Week 9)
negotiation_teacher_help	Student seeks help from teacher, or teacher proactively intervenes to pro-	Student raises hand, calls teacher, or teacher proac-	Teacher provides whole-class instruction (not targeting specific	When S7 gets stuck in Week 5, teacher approaches and

H.2.2 Output Modification Codes

These codes are used to label students' adjustments to their language output during interaction with the AI, corresponding to the "noticing" and "hypothesis-testing" functions of the Output Hypothesis (Swain, 2005).

Table 8 Table H.2.2 Coding scheme for output modification

Code Name	Definition	Inclusion Criteria	Exclusion Criteria	Data Example
output_noticing	Student demonstrates awareness of their own language problem, usually through self-commentary, pausing, or self-correction	Student explicitly says "Did I say it wrong?" "It should be..." etc.; student shows noticeable self-correction pause	Student simply repeats original utterance without showing awareness of error	"Did I say 'right' not clearly enough?" (S3, Week 6 interview)
output_hypothesis_testing	Student tests their hypothesis about language form through output, usually by trying a new expression and observing AI's response	Student tries new expression; student explicitly states in interview "I wanted to see if the AI would understand it this way"	Student simply repeats original utterance; student uses already confirmed correct expression	"I was thinking about whether to use 'le'... The AI understood, so it must be right." (S11, Week 18 interview)
output_modification_lexical	Student modifies vocabulary in output	Student substitutes, adds, or deletes one or more words; sentence structure remains basically unchanged	Student modifies grammatical structure (code as output_modification_syntactic)	Changes "My father tall" to "My father is very tall" (adding "very")
output_modification_syntactic	Student modifies grammatical structure or word order in output	Student changes sentence structure, word order, or adds/deletes grammatical markers (e.g., aspect marker "le", disposal marker "ba")	Student only modifies vocabulary (code as output_modification_lexical)	Changes "Saturday I play ball do homework watch TV" to "Saturday I first play ball, then do homework, finally watch TV" (adding connectives)
output_metalinguistic_reflection	Student engages in metalinguistic reflection about the output process, explicitly expressed in interviews or on task cards	Student describes their thought process in interviews or task cards; student uses metacognitive expressions like "The process of teaching it is also learning"	Student only describes behavior without reflection	"The process of teaching it is also learning." (S7, Week 18 interview)

H.2.3 Affective Expression Codes

These codes are used to label students' emotional states during interaction with the AI, corresponding to the Affective Filter Hypothesis (Krashen, 1982).

Table 9 Table H.2.3 Coding scheme for Affective expression

Code Name	Definition	Inclusion Criteria	Exclusion Criteria	Data Example
affective_anxiety_high	Student shows obvious nervousness, unease, or anxiety	Student's voice becomes softer, head lowers, silences, body stiffens, or self-reports nervousness	Student is merely focused or serious (without accompanying tense body language)	S7 lowers head, voice too soft to hear, silent for 5 seconds in Week 1
affective_anxiety_low	Student shows relaxed, at ease state	Student's voice is normal or loud, body relaxed, or self-reports not nervous	Student shows excitement (code as affective_positive_excitement)	"Speaking to the AI, if you say it wrong it won't laugh at me." (S7, Week 6 interview)
affective_positive_excitement	Student shows excitement, joy, or enjoyment	Student laughs, cheers, shares joy with peer, or self-reports "It's really fun"	Student merely completes task calmly (code as affective_neutral)	S7 smiles and says to peer "It's so stupid" (Week 5)
affective_positive_satisfaction	Student shows calm satisfaction, distinct from excitement	Student smiles, nods, or self-reports "It's helpful but not exciting"	Student shows excitement (code as affective_positive_excitement); student shows frustration (code as affective_frustration)	S3 nods and says "Right" after AI understands (Week 13)
affective_frustration	Student shows frustration, impatience, or annoyance	Student sighs, speeds up speech, frowns, or self-reports "It's so annoying"	Student shows anxiety (code as affective_anxiety_high)	S11 sighs and says "You got it wrong again" (Week 5); S13 says "It finally understood, took three tries" in a tired tone (Week 9)
affective_neutral	Student shows neutral emotional state, with no obvious positive or negative emotion	Student's expression is calm, tone is flat, no language expressing emotion	Student shows any identifiable positive or negative emotion	S11 maintains calm expression throughout activity, makes no comment after completing task (Week 5)

H.2.4 AI Role Perception Codes

These codes are used to label the roles students assign to the AI, corresponding to Luckin et al.'s (2016) AI role framework and extensions from this study.

Table 10 Table H.2.4 Coding scheme for AI role perception

Code Name	Definition	Inclusion Criteria	Exclusion Criteria	Data Example
role_tool	Student perceives AI as a passive, controllable tool, similar to a dictionary or voice input device	Student uses metaphors like "tool," "dictionary," "calculator"; student shows instrumental use posture (no emotional investment, task-oriented)	Student anthropomorphizes AI (e.g., gives it a name, discusses AI with peers)	"It's just a tool. Like Google Translate." (S11, Week 6 interview)
role_peer_need_help	Student perceives AI as a peer with limited ability who needs to be taught	Student uses metaphors like "little kid," "foreigner just learning Chinese"; student shows posture of "teaching the AI"	Student perceives AI as opponent or game object (code as role_opponent)	"It's like a little kid. It doesn't understand anything. You have to teach it." (S3, Week 6 interview)
role_opponent	Student perceives AI as something that can be "defeated," enjoying the process of correcting it	Student uses metaphors like "opponent," "game"; student shows competitive posture; student enjoys AI's mistakes	Student merely calmly corrects AI (code as role_peer_need_help or role_tool)	"It's like an opponent in a game. You have to beat it." (S7, Week 6 interview)
role_coach	Student perceives AI as a role that provides prompts, guides learning, but does not directly give answers	Student uses metaphors like "coach," "guide"; student describes AI "pushing me to think"	Student perceives AI as teacher who directly gives answers (code as role_teacher)	"Sometimes it asks 'And then?' That's its way of prompting you to continue." (S11, Week 18 interview)
role_teacher	Student perceives AI as an evaluator, corrector, authority figure	Student uses metaphors like "teacher," "exam"; student shows posture of fearing being judged by AI	Student merely accepts AI's feedback (may still be tool role)	"At first I thought it was like a teacher. Because it tells you if you're right or wrong. I was a bit afraid of it." (S7, Week 6 interview)
role_game_partner	Student perceives AI as an entertainment object to joke about and share with peers	Student discusses AI's "stupidity" with peers; student treats AI's mistakes as something to laugh about	Student enjoys correcting AI alone (code as role_opponent)	S7 says to peer "It's so stupid" and both laugh (Week 5)
role_fluid	Student assigns different roles to AI within the same time period, or role perception changes	Student shows two or more role perceptions within the same observation period; role per-	Student's role perception is stable (code under other single role codes)	"Sometimes it's like a tool, sometimes it's like a little kid that needs

H.2.5 Privacy and Ethics Codes

These codes are used to label data related to privacy protection and ethical decision-making, corresponding to RQ3.

Table 11 Table H.2.5 Coding scheme for privacy and ethics

Code Name	Definition	Inclusion Criteria	Exclusion Criteria	Data Example
privacy_teacher_action	Privacy protection measures or ethical decisions taken by the teacher	Teacher implements privacy protection behaviors (reminders, clearing records, deleting data); teacher makes ethics-related decisions	Teacher's general teaching decisions (unrelated to privacy)	Researcher reminds entire class "Don't type your name" (Week 2 reflection log)
privacy_student_behavior	Privacy awareness or related behaviors shown by students	Student proactively asks privacy questions; student protects own or others' privacy; student unintentionally leaks information	Student's general AI use behaviors	S8 unintentionally types full name (Week 4); student asks "Will the AI remember what I said?" (Week 1)
privacy_parent_concern	Privacy or ethics-related concerns expressed by parents	Parent expresses concerns about data privacy, screen time, technology dependence, etc., in questionnaire or interview	Parent expresses other opinions about course content (unrelated to privacy)	"My biggest concern is my child's voice being recorded." (P4, Week 1 questionnaire)
privacy_parent_acceptance	Acceptance or attitude change expressed by parents	Parent states in questionnaire or interview "I'm less worried now," "I support using it," etc.	Parent has never expressed concerns (not applicable)	"I'm less worried now. I trust that you are paying attention." (P7, Week 18 interview)
ethics_dilemma	Ethical dilemmas unexpected by the teacher and the process of handling them	Teacher records unexpected ethical incidents; teacher describes how dilemma was handled	Routine privacy protection measures (code as privacy_teacher_action)	AI generates inappropriate content, teacher intervenes (Week 10 reflection log)

H.2.6 Teacher Reflection Codes

These codes are used to label content in teacher reflection logs, corresponding to the researcher's self-reflection process.

Table 12 Table H.2.6 Coding scheme for teacher reflection

Code Name	Definition	Inclusion Criteria	Exclusion Criteria	Data Example
reflexivity_emotion	Teacher's reflexive reflection on their own emotional state	Teacher records their emotional reactions; teacher reflects on how emotions might affect the research	Teacher records student emotions (code under affective expression codes)	"I was very happy during class today... But I also realize that my 'happiness' might affect my observations." (Week 5 reflection log)
reflexivity_bias	Teacher's reflection on their own potential biases	Teacher acknowledges possible biases; teacher describes strategies to address biases	Teacher records objective facts (without reflective component)	"I noticed I paid more attention to S7 and less to S11... I need to remind myself." (Week 1 reflection log)
teaching_decision	Teacher records teaching decisions and the rationale behind them	Teacher describes unplanned adjustments; teacher explains why a particular decision was made	Teacher describes original plan (without adjustment)	"Instead of cutting students' conversation time, I shortened the final 'presentation' segment. I think conversation practice is more important." (Week 1 reflection log)
action_plan	Teacher records action plan for next lesson	Teacher explicitly writes "Next time I will..."	Teacher's general wishes (without specific actions)	"Starting next week, I need to give a 10-second reminder at the beginning of each Al lesson: 'Don't type your name.'" (Week 1 reflection log)

H.3 Examples of Coding Application

The following examples show how codes are applied to raw data segments.

Example 1: Classroom Observation Record (Week 5, S7)

Raw Data:

S7 says to AI: "Go straight first, then turn left."
AI: "You want to go to the park? Turn left first?"
S7 immediately responds: "No! Not left turn. Go straight first!" He uses his hand to gesture a pushing forward motion. Then he turns to his peer, smiles, and says: "It's so stupid."

Coding:

Data Segment	**Code**
"No! Not left turn. Go straight first!"	negotiation_correction
Uses hand to gesture pushing forward motion	negotiation_gesture
Turns to peer, smiles, and says "It's so stupid"	role_game_partner
Smiles	affective_positive_excitement

Example 2: Student Interview (S7, Week 18)

Raw Data:

"The process of teaching it is also learning. Because you have to think about how to explain it so it can understand, you have to think about how to say that word, or use a different way of saying it. It's the same as learning yourself."

Coding:

Data Segment	**Code**
"The process of teaching it is also learning"	output_metalinguistic_reflection
"You have to think about how to explain it so it can understand"	output_noticing
"Use a different way of saying it"	negotiation_paraphrase

Example 3: Teacher Reflection Log (Week 1)

Raw Data:

"To be honest, I was a bit nervous before class today this was my first time using Al for a formal activity in a real classroom. I was afraid of technical problems and afraid students wouldn't like it. After watching the recording, I noticed that I paid more attention to S7 and less to S11. This might be because S7's problems were more 'visible.' I need to remind myself not to focus only on students who are 'having problems."'

Coding:

Data Segment	Code
"I was a bit nervous before class today"	reflexivity_emotion
"I noticed that I paid more attention to S7 and less to S11"	reflexivity_bias
"I need to remind myself"	action_plan

Example 4: Parent Interview (P7, Week 18)
Raw Data:
"At the beginning of the semester, I was worried about my child's conversations being recorded. But during the semester, I observed and also asked my child about it several times. He said the teacher reminds them to close the window every time. I also saw that you don't ask students to use their real names. So I'm less worried now."
Coding:

Data Segment	Code
"At the beginning of the semester, I was worried about my child's conversations being recorded"	privacy_parent_concern
"I'm less worried now"	privacy_parent_acceptance

H.4 Coding Usage Instructions

H.4.1 Coding Principles

Principle	Description
One segment can have multiple codes	The same data segment may correspond to multiple codes (e.g., S7's "It's so stupid" is coded for both role and affect)
Code mutual exclusivity	Within the same dimension, the same segment should not have two mutually exclusive codes (e.g., the same affective segment cannot be coded as both affective_positive_excitement and affective_frustration)
Codes require evidence	Each code must have clear data evidence to support it; avoid "over-coding"
Codes can be revised	The coding scheme is dynamic; codes can be added, deleted, or modified during analysis based on new data

H.4.2 Coding Reliability Assurance

To ensure consistency and trustworthiness of coding, this study adopted the following measures:

1. **Intra-rater consistency:** The researcher recoded the same data (approximately 20% of the sample) after a one-month interval and compared consistency between the two coding sessions. Consistency reached above 85%.
2. **Peer review:** An independent peer researcher who did not participate in this study independently coded one focal student's interview transcript (approximately 10% of the sample), compared coding differences, discussed discrepancies, and reached consensus.
3. **Coding log:** The researcher recorded decisions and questions during the coding process, forming a "coding decision log" for traceability and review.

Note: All student names are pseudonyms. Some example responses are composites for illustration.

Appendix I. Informed Consent Forms

I.1 Student Informed Consent Form

Research Project: AI-Assisted Chinese Learning

Hello!

I am [name], your Chinese teacher. I am doing a research project to understand whether AI (Artificial Intelligence) can help us learn Chinese. I would like to invite you to participate in this research.

What is this research about?

- We will use some AI tools (like ChatGPT) in Chinese class
- I want to see how you learn when you talk to the AI
- I will also ask you some questions to hear your thoughts

What do I need to do if I participate?

- Attend Chinese class as usual and talk to the AI
- Sometimes I will record your screen when you talk to the AI (no face recording)
- Sometimes I will ask you a few questions (about 10-15 minutes)

What are the benefits of participating?

- You may have more opportunities to practice speaking Chinese
- Your ideas will help me improve my teaching

What are the risks of participating?

- There are no major risks. If you feel uncomfortable, you can stop at any time.

Your rights:

- You decide whether to participate or not
- If you want to stop, you can say "I don't want to participate anymore" at any time
- Not participating will not affect your Chinese grade
- I will keep everything confidential. Your name will not appear in the research paper.

Do you want to participate?

☐ I want to participate in this research
☐ I do not want to participate

Your name: _____

Date: _____

Thank you! If you have any questions, you can ask me anytime.

1.2 Parent Informed Consent Form

Research Project: A Qualitative Study of AI-Assisted Chinese as a Second Language Learning in Middle School

Researcher Information

Item	**Content**
Researcher Name	[Name]
Position	Chinese Teacher
Institution	[School Name]
Contact	[Email] / [Phone]

Dear Parent/Guardian,

Your child is invited to participate in a research study on "AI-Assisted Chinese as a Second Language Learning in Middle School." Before you decide whether to allow your child to participate, please read the following information carefully. If you have any questions, please feel free to contact me.

I. Purpose of the Study

The purpose of this study is to understand how Artificial Intelligence (AI) tools (such as ChatGPT) are used in middle school Chinese as a second language classrooms, and how students interact with and perceive the AI. Specifically, this study aims to answer the following questions:

1. How do students engage in conversation and negotiate meaning with the AI?
2. How do students perceive the role of AI in their Chinese learning?
3. How does the teacher address privacy and ethical issues in AI use?

The findings of this study will help improve the design and practice of AI-assisted language teaching.

II. Study Procedures

If you agree to allow your child to participate, here is what your child will be asked to do:

Activity	Frequency/Duration	Description
Regular participation in AI-assisted Chinese lessons	Once per week, 50 minutes each	Your child will use AI tools (ChatGPT, Duolingo, Canva, etc.) for Chinese conversation practice in class
Classroom video recording (panoramic view only, no close-up faces)	Once per week	A camera will be placed at the back of the classroom to record the panoramic view. Recordings are for research analysis only and will not be made public
Screen recording	Once per week	Recording of your child's conversation with the AI (screen and voice only, no face)
Task card completion	Once per week	Brief written task card completed in class (about 5 minutes)
Interview participation	Once mid-semester + once end of semester, 20-25 minutes each	One-on-one interview discussing your child's feelings and views about the AI. Interviews will be audio-recorded
Parent questionnaire/interview	Once beginning + once end of semester	You will be invited to complete an open-ended questionnaire (about 10 minutes) or participate in an interview (15-20 minutes)

Duration of the study: One semester (approximately 18 weeks)

III. Risks and Discomforts

This study poses minimal risk to your child. Possible discomforts may include:

- Occasional frustration when interacting with the AI (e.g., when the AI does not understand)
- Mild self-consciousness when being observed

Mitigation measures:

- The teacher will closely monitor students' emotional states and provide support when needed
- Students may stop participation at any time without giving any reason
- Classroom observations are conducted in a natural teaching environment to minimize disruption

IV. Benefits

Direct benefits your child may receive from participating in this study include:

- Increased opportunities to practice speaking Chinese
- Experience of success in speaking Chinese in a low-anxiety environment
- Development of AI literacy (an important 21st-century skill)

Indirect benefits of this study include:

- Helping the teacher improve the design of AI-assisted teaching
- Providing experience that other schools and teachers can learn from

Your child will not receive any financial compensation for participating in this study.

V. Confidentiality and Privacy

The following measures will be taken to protect your and your child's privacy:

Data Type	Protection Measures
Student names	Pseudonyms (e.g., S1, S2) will be used; real names stored only in encrypted files
Classroom video recordings	Panoramic view only, no close-up faces; viewed only by the researcher; destroyed 12 months after study completion
Screen recordings	Only Al conversation content recorded, no face; deleted immediately after transcription
Interview audio recordings	Deleted immediately after transcription; all identifiable information removed from transcripts
Task cards	Scanned and stored in encrypted folders; paper copies destroyed after scanning
Parent questionnaires	Anonymous, no name required

Data usage scope: All data will be used for this study only and not for any other purpose. No identifiable personal information will appear in any published research findings.

Data storage: All electronic data will be stored on an encrypted hard drive; paper documents will be kept in a locked cabinet. Only the researcher will have access.

VI. Voluntary Participation and Withdrawal

- Your and your child's participation is completely voluntary
- You may choose to agree or decline; declining will not affect your child's Chinese grade or relationship with the teacher
- Your child may withdraw from the study at any time without giving any reason
- If you or your child decides to withdraw, please notify the researcher. All data collected about your child will be deleted

VII. Informed Consent

I have read and understood the information above. I have had the opportunity to ask questions, and my questions have been answered satisfactorily.

Please check the following box:

☐ I agree to allow my child to participate in this study
☐ I do not agree to allow my child to participate in this study

If you agree, please complete the following information:

Item	Content
Child's Name	
Parent's Name	
Parent's Signature	
Date	

Contact information (if you wish to receive a summary of the research findings):
Email: _____ **or Phone:** _____

Researcher's Signature
I confirm that I have explained the nature, purpose, risks, and benefits of the study to the participant/parent.
Researcher's Signature: _____
Date: _____

Please return the signed consent form to the researcher. Thank you for your trust and support!